

Breaking down the barriers

Many Japanese researchers are concerned that they don't compete on a level playing field when it comes to international science. Language and cultural barriers may be partly to blame. But the perception is more forbidding than the reality.

Science is supposed to be universal, crossing borders, overcoming linguistic barriers, and piercing through cultural differences. But Japanese researchers have long fretted that they don't compete effectively on the international scene — and policy-makers have racked their brains to find ways of overcoming the problem.

The latest initiative is a government-backed plan to publish a series of online journals designed to showcase top-quality Japanese research (see page 868). It is, in large part, a reaction to a concern among Japanese scientists that leading journals may be biased against them.

Such feelings are by no means unique to Japan. Speak to researchers elsewhere in Asia and across much of Europe and you will hear similar complaints. In a world of science in which English is the lingua franca, it's not surprising that non-native speakers sometimes feel hard done by. When your paper is rejected without even being sent for peer review, it's tempting to wonder whether a manuscript written in more polished English would have been given the same treatment.

Concerns about language barriers extend even to the élite of Japanese research. Kenichi Fukui of Kyoto University won the Nobel Prize in Chemistry in 1981 for his theoretical insights into chemical reactions. But even after this triumph, Fukui had to deal with critics who disliked his 'inelegant' English formulation of his ideas.

Such is the dominance of English, however, that it pervades most attempts to improve the competitiveness of Japanese science. Even the proposed online journals will be published in English, for example. Institutes established to give Japan a world-class presence in 'hot' fields — such as the Brain Science Institute in Saitama and the Center for Developmental Biology in Kobe — conduct Anglophone lab meetings. There is also a plan to establish an English-speaking graduate school on the poor southern island of Okinawa (see *Nature* **416**, 774; 2002). And the desirability of making funding applications in English to facilitate international evaluation is surfacing in current discussions over the reform of Japan's university system (see page 875).

Lab visits

Many individual Japanese scientists, of course, have managed to compete successfully on the international stage, and have done so without worrying too much about whether language barriers will hold them back. They attend conferences, interacting in imperfect English if need be. They invite foreign researchers to visit their laboratories. When they need to visit a foreign lab, they do so, even though they might struggle to converse. And they ask foreign colleagues to comment and criticize their draft manuscripts — both for their scientific reasoning and the standard of their English. Some even hire English-speaking editors to polish their papers before submission.

A little English, it seems, can go a long way. It is a forgiving tongue, in which precise grammar is often not necessary for the meaning to get across. Unfortunately, however, many Japanese scientists seem to be shy of engaging with foreign peers, fearful that

they will be let down by their grasp of English. The way in which the language is taught in Japanese schools may be partly to blame — the emphasis is on memorization and grammar, rather than expression and communication. As a result, almost all Japanese have a fairly broad English vocabulary, but many lack the confidence to apply it in conversational situations.

So what about existing journals? Accusations of bias — which may be unconscious — are difficult to address. But bear in mind that leading journals reject the vast majority of papers that they receive, so there's no reason to jump to the assumption that a rejection has anything to do with the author's nationality or ability to write in English. *Nature* has expanded its pool of Japanese reviewers, and hopes to continue to do so. Our editors — more than a quarter of whom have English as their second language — are happy to work with authors to improve a paper's standard of written English to bring out important scientific points.

Opening up

Specific initiatives to improve the international standing of Japanese science may have some merit, but aren't without difficulties. The proposed new online journals will face a tough task convincing scientists to submit their best work to them instead of the likes of *Nature*, *Science*, *Cell*, *Physical Review Letters* and the *Journal of the American Chemical Society*. The new publications will also have to garner readers and reviewers from around the world if they are to have a reasonable chance of success.

Both the Brain Science Institute and the Center for Developmental Biology, meanwhile, have set and then struggled to meet quotas for foreign staff. Generous grant programmes designed to bring foreigners to Japan are undersubscribed. And there is widespread scepticism about whether the Okinawa graduate school will succeed in its plan to recruit half of its professors from abroad.

Given the difficulty of creating islands of international excellence within the Japanese system, what more can be done to make the system as a whole more international in its outlook? Opening Japanese funding opportunities to applicants writing in English would certainly be a step in the right direction.

But officialdom isn't the only thing that has to change. Postdocs and other researchers visiting Japan often complain that they had problems communicating with their supervisor or with colleagues — problems that have little to do with specific linguistic barriers. Rather, they feel ignored and deprived of opportunities to discuss their research. The reasons are unclear, but perhaps shyness among Japanese scientists about communicating with foreigners could again be a factor.

Internationalization is a two-way street. Japanese labs must become more open and welcoming places for researchers coming from abroad, and visiting foreigners must be prepared to adapt to their new environment (see page 877). In the end, the internationalization of Japanese science is in the hands of individual scientists and their ability to get around language or other supposed barriers. More often than not, the perception of a linguistic or cultural barrier is more inhibiting than the barrier itself. ■



Tribes query motives of knowledge databases

Rex Dalton, Kelowna

Attempts to set up databases of traditional knowledge are coming under attack from the very groups they are intended to benefit.

The World Intellectual Property Organization (WIPO), the Geneva-based body that promotes intellectual property rights, is keen to establish databases in which indigenous groups would record their cultural knowledge. Patent examiners would be encouraged to check a country's databases to see if a new idea, such as a plant-based medicine, is part of the traditional knowledge of that nation.

But delegates at the World's Indigenous Peoples conference, held on 16–19 October in Kelowna, British Columbia, described how some groups are concerned that the databases could be used to exploit their cultural heritage. Some Canadian tribes, for example, already have individual databases. But a national project, conducted on a shoe-string budget by the Centre for Traditional Knowledge at the Canadian Museum of Nature in Ottawa, has suffered because indigenous groups are reluctant to make their databases available. The groups fear that secret traditions associated with plants or substances will be stolen, according to government officials working on the project.

Verna Miller, a custodian of traditional knowledge for the Nicola Tribal Association in Merritt, British Columbia, adds that her



tribe guards its secrets so closely that even she doesn't know the password to its database — only the tribal elders have it.

Overcoming such diffidence may be difficult. Advocates for indigenous groups fear that the databases could be used by Western companies and scientists aiming to profit from traditional knowledge. Controversial cases, such as the award and subsequent retraction of a US patent for a turmeric-based wound remedy used in India (see *Nature* 389, 6; 1997), have generated widespread mistrust of patent organizations such as WIPO. "I can't think of a more inappropriate place for discussions on traditional

knowledge," says Debra Harry, executive director of the Indigenous People's Council on Biocolonialism in Wadsworth, Nevada.

WIPO officials say they hope that involving indigenous groups in the development of the databases should help. The organization is consulting over 50 developed and developing countries about their requirement for databases and has set up a web portal for patent examiners, which already links to established databases in India and China. Two indigenous groups have also been granted observer status at meetings of the WIPO committee on traditional knowledge. "We want to make sure people are confident they have input to the process," says Tony Taubman, head of the organization's traditional knowledge division. "It would be fatuous to sit in Geneva and dream up a one-size-fits-all solution."

WIPO will next year publish reports on the databases and on other issues, which it hopes will clarify how traditional knowledge can be protected. But delegates at the Kelowna meeting are yet to be convinced that the idea will work. "I am highly sceptical that a national database will fly," says a Canadian scientist who works closely with indigenous groups. "At the heart of the debate is control of the knowledge. We have a long way to go. This takes commitment I haven't seen at a provincial or national level."

♦ www.wipo.org

Low stocks prompt calls for North Atlantic fishing ban

Quirin Schiermeier, Munich

Unprecedented restrictions on commercial fishing in European Union (EU) waters have been proposed by leading fisheries scientists.

The Copenhagen-based International Council for the Exploration of the Sea (ICES), which provides scientific advice to fisheries-management authorities, says that all cod and haddock fisheries in the North Sea, the Skagerrak, the Irish Sea and off the western coast of Scotland should be closed for an undetermined period to allow stocks to recover.

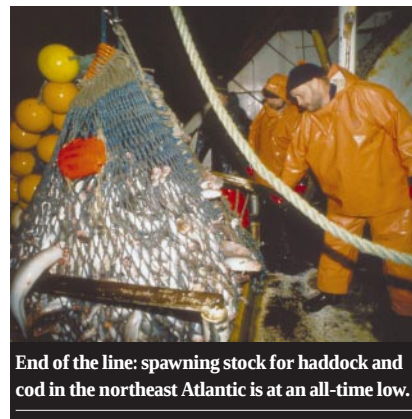
The proposal has sent shock waves through the EU fishing industry, which would be in danger of substantial economic losses if prime fishing areas were closed. But scientists say that only drastic measures will save the stocks. According to the ICES assessment, carried out using data from fisheries laboratories in the council's 19 member states, the 'spawning stock biomass' (SSB) of cod and haddock in the northeast Atlantic are at new historic lows. The number of young fish surviving to adulthood is also the lowest on record.

Closure of parts of the most heavily fished oceans has long been espoused by conservation biologists as one way to rebuild stocks (see *Nature* 419, 664; 2002). EU member states often relax the stringent catch quotas recommended by ICES in order to appease their fishing industries, but the scale of the problem means that politicians must now accept large short-term losses, says Hans Lassen, a marine biologist who coordinates the work of groups that contribute to ICES reports.

Fisheries scientists also warn that the relatively stable cod and haddock catches in the northeast Atlantic in recent years are not reliable indicators of the abundance of fish in EU waters. They point to the Canadian Grand Banks fisheries off Newfoundland, which were closed in 1992 after the cod population collapsed. Canadian authorities failed to realize that intensified effort by fishing vessels was masking a decline in stocks. Only independent data gathered by research vessels can give an unbiased picture of the state of the stocks, the ICES report argues.

"It is an inescapable conclusion that the cod stocks around the United Kingdom are seriously depleted," says Robin Cook, chief executive of the Fisheries Research Services' Marine Laboratory in Aberdeen, UK. "Given that all management practices tried before have failed, closures are about all you can do in this rather extreme situation."

♦ www.ices.dk



End of the line: spawning stock for haddock and cod in the northeast Atlantic is at an all-time low.

T. RAUPACH/STILL PICTURES



Bloody battle: the mosquito *Aedes aegypti* remains a scourge of health officials fighting dengue fever.

Mosquito researchers deny plotting secret biowarfare test

Kendall Powell, Washington
and K. S. Jayaraman, New Delhi

A recent admission by the United States that it conducted a biological-warfare test using mosquitoes in 1965 has reopened old wounds over a much larger, but aborted, mosquito research project in India.

The Indian project was run by the World Health Organization (WHO) and the Indian Council of Medical Research (ICMR), but received funding from the US government. Researchers involved had planned to release hundreds of thousands of sterile male *Aedes aegypti* mosquitoes in the town of Sonipat, 100 kilometres north of New Delhi. But the project was cancelled by the Indian government in 1975 after Indian researchers claimed that its real purpose was to study the logistics of using yellow fever as a biowarfare agent (see *Nature* **251**, 177–178; 1974). The disease does not exist in India, but is transmitted by *A. aegypti* and was considered at the time to be a potential biowarfare agent.

The allegations have always been strenuously denied by the European and US researchers involved. They insist that their aim was to eradicate the mosquito population, which transmits dangerous diseases such as dengue fever. But Indian scientists say that new documents released by the US Department of Defense show that the United States was conducting similar biowarfare experiments elsewhere. The documents, released on 9 October, list 27 secret chemical and biological tests conducted at the height of the Cold War. One of the tests involved releasing *A. aegypti* mosquitoes off the coast of the uninhabited Baker Island in the South Pacific to track the logistics of a mosquito-borne viral attack.

“The latest revelation that the Baker Island release was indeed a biowarfare experiment vindicates the closure of the US project in India,” says N. P. Gupta, an ICMR member who was director of the National Institute of

Virology in Pune when the project was cancelled. P. K. Rajagopalan, a retired entomologist who worked on the Sonipat project for the ICMR, points out that both programmes aimed to track the dispersal patterns of marked mosquitoes, and are similar enough to confirm government suspicions.

But researchers outside India vehemently deny that the Sonipat project was anything other than a legitimate public-health research project. “It was a very important species of mosquito to try to get rid of,” says Scott Halstead, a dengue expert who served as a WHO consultant for the project. “Dengue and dengue haemorrhagic fever are much more of a problem in India now. Here we are 30 years later and we still don’t have a way of dealing with the mosquito — the mosquito is winning.”

Others say that the only similarity between the two projects was the species of mosquito used, and that the data gathered in Sonipat would not have been useful for biowarfare purposes. In the Baker Island test, which involved female mosquitoes only, researchers tracked how many mosquitoes reached traps on the island, and recorded the number of bites volunteers on the island received. “We went to a lot of trouble to make sure that we were releasing 99% sterile males,” points out Chris Curtis, a WHO medical entomologist on the project who is now at the London School of Hygiene and Tropical Medicine. Unlike female mosquitoes, males do not bite or transmit viruses, he explains. “We couldn’t have possibly produced useful data for biowarfare.”

“Just because the words mosquito and United States can be threaded into one sentence does not implicate the project in India,” adds Halstead, now at the Uniformed Services University of the Health Sciences in Bethesda, Maryland. “That project was only biological warfare from the perspective of the mosquito.” ■

Penicillin paper restores Fleming’s healthy reputation

Tom Clarke, London

Inspired by musicologists’ use of fragmented scores to complete the unfinished works of great composers, a British researcher has pieced together Alexander Fleming’s laboratory scribbles to recreate a paper that he says restores the reputation of the much-maligned discoverer of penicillin.

Fleming published details of the antibiotic effects of a mould that had killed off bacterial cultures in his poorly sterilized petri dishes, but never isolated penicillin from the mould or published work on its potential as a drug. Many of Fleming’s contemporaries and biographers have accused him of being messy and lazy, and of losing interest in his chance 1928 discovery, even though he went on to take much of the credit for discovering the first antibiotic drug.

“I hope my version of this paper will once and for all scotch the idea that Fleming was some idle dilettante who did little to develop what is arguably the most important drug in medicine,” says Milton Wainwright, the University of Sheffield microbiologist who has written the paper that Fleming could have produced, but didn’t (*M. Wainwright Perspectives in Biology and Medicine*, in the press).

The posthumous paper, which is based on Wainwright’s studies of Fleming’s lab notes and personal communications, describes the antibacterial properties of other airborne moulds. Wainwright emulated Fleming’s writing style and prepared the paper in a format suitable for the now defunct *British Journal of Experimental Pathology*, where Fleming published his first work on penicillin.

Wainwright also argues that Fleming fully realized penicillin’s potential prior to 1940. It was around this time that biochemists Howard Florey and Ernst Chain at the University of Oxford first isolated penicillin and demonstrated its potential as a drug. Fleming perhaps delayed publication while he gathered data for a *magnum opus* on penicillin, argues Wainwright.

Composer and musicologist Anthony Payne, who has used the records kept by Edward Elgar to complete the British composer’s *Symphony No. 3*, says that there are parallels between the two works — before the third symphony was performed, scholars regarded Elgar as an artist in decline. “My work proved that what people had been saying about him was completely wrong,” says Payne. ■

Biologists join drive to turn down the lights

Steve Nadis, Boston

The fight against light pollution is becoming a multidisciplinary endeavour. Last week's meeting of the International Dark-Sky Association (IDA), which would once have been an event purely for astronomers, featured researchers working on topics as diverse as migratory birds and breast cancer.

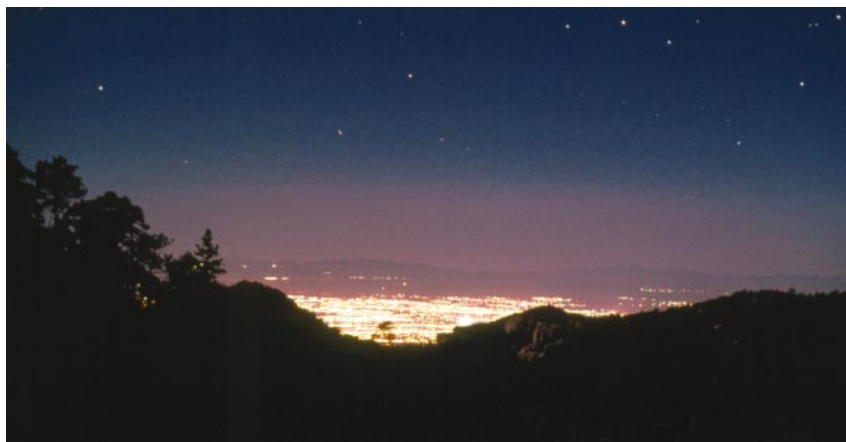
Astronomers attracted little interest when they first complained about the glow from urban areas in the 1950s, but the dark-sky movement has grown rapidly since the IDA was formed in 1988. Astronomy was deliberately played down at this year's meeting, and was the subject of only one of some 40 presentations. "If it's just a concern for astronomers, many people will dismiss this as a narrow, special-interest group," says Dan Green of the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, one of the event's organizers. "We're stressing the security, energy-efficiency, health and ecological benefits of good lighting."

Cancer epidemiologist Richard Stevens, from the University of Connecticut in Farmington, provided one example of this wider focus. He discussed potential links between the increased incidence of breast cancer, especially in the industrialized world, and the way in which exposure to light at night can disrupt natural secretion cycles of the hormone melatonin. Low levels of melatonin, which is thought to be responsible for our daily, or circadian, biological rhythms, have been linked to breast cancer.

Stevens cited recent studies of visually impaired men and women, which show that breast-cancer risk declines with decreasing vision (P. K. Verkasalo *et al. Br. J. Cancer* **80**, 1459–1460; 1999). Although Stevens' work is not specifically concerned with light pollution, other meeting participants tried to make the link, suggesting that 'light trespass' into homes and apartment buildings from street lights and other sources is more than capable of perturbing circadian rhythms.

Other speakers addressed risks posed to migrating birds by outdoor artificial lighting. Ornithologist Richard Podolsky of the Technology Planning and Management Corporation, an environmental services organization based in Scituate, Massachusetts, described his research of offshore birds. Atlantic puffins, which rely on the Moon for navigation, become confused by illuminated structures at sea, says Podolsky, often circling them until they crash or die of exhaustion.

Delegates at the meeting also paid tribute to the progress made by the IDA. David Crawford, an astronomer from Kitt Peak National Observatory in Tucson, Arizona, and the IDA's executive director and co-founder, began the campaign in the late 1960s by joining the Illuminating Engineer-



Ray ban: astronomers have long campaigned for limits to the amount of light pollution from cities.

ing Society of North America (IESNA), an organization that aims to set standards for outdoor lighting. "Two decades ago, we hated those guys and used to run from them," recalls IESNA president Randy Reid. "We

thought they were a bunch of left-wing lunatics. But Crawford persuaded us that he didn't want to turn the lights out, he just wanted us to direct light where it was needed, without illuminating the whole sky."

Japan plans web of English journals

David Cyranoski, Tokyo

Japan is aiming to increase its already significant presence in international research by launching a set of electronic English-language journals.

Last month the education ministry requested ¥180 million (US\$1.4 million) from the government, which is expected to be granted by the end of the year, to launch the plan.

The National Institute of Informatics (NII) in Tokyo will provide the technology required for the journals. The institute will also select the fields in which the first four or five journals will be launched. The relevant Japanese academic societies will recommend editorial-board members. Each journal will ask referees from across the world to review its papers, and will be expected to be self-supporting after three years.

The NII will collaborate with the Scholarly Publishing and Academic Resources Coalition (SPARC), an initiative established by libraries in the United States and Europe to reduce journal costs by distributing online publications at affordable prices (see *Nature* **398**, 272; 1999). SPARC will also help the NII to develop a business model for the journals, and will help it to find buyers for the new online journals, say planners at the education ministry.

If the new online journals provide high impact and value for money, libraries will favour them over more expensive ones, says an

official from one national university library. But the plan is not just about cost-cutting. "The project came about also because the education ministry is keen to enhance the reputation of Japan's research," says Syun Tutiya, a philosopher and former librarian at Chiba University, north of Tokyo, who hopes to set up an equivalent to SPARC in Japan.

Ministry officials are playing down the implication that European and US journals neglect Japanese research, but say that researchers are frustrated at having to submit to US and European journals. "Many researchers prefer to submit their papers to a Japanese journal, and will continue to do so, because they're concerned they won't get fair treatment elsewhere," says Mikiko Tanifuji of the Institute of Pure and Applied Physics in Tokyo, which publishes four journals, including the *Japanese Journal of Applied Physics*.

Japanese scientists say that a successful journal could be a boon for research if it is able to bring in submissions and peer reviewers from around the world. But they worry that the nominally international effort could become a nationalistic enterprise that merely advertises Japanese research.

The key to the journals' success will be finding fields in which Japan excels, researchers say. "The question is whether the journals will actually be used," says Tanifuji.

♦ www.nii.ac.jp

MIT gets plugged in for global data archive

Declan Butler

Plans by the Massachusetts Institute of Technology (MIT) look set to transform the way in which academics publish and archive their results and raw data.

On 4 November, the institute will launch its DSpace electronic archive, a joint venture with technology firm Hewlett-Packard of Palo Alto, California.

The ultimate goal of the system is to capture MIT's entire intellectual output in a stable archive, and to extend it to create a seamless global network of similar archives at other research institutions. These multiple databases could be searched as if they were a single entity, and specialized collections could be built by drawing data from participating archives.

More than 40 academic institutions worldwide are considering adopting DSpace, and seven universities, including the University of Cambridge in Britain, are installing and evaluating the system. The move towards such archives is in part driven by libraries, which are looking for cheaper alternatives to costly academic journals.

The technology involved is not rocket science, admits MacKenzie Smith, associate

director for technology at MIT Libraries and head of the DSpace project. The main innovation is in transforming the often anarchic sprawl of researchers' websites into a professionally managed system, she says. Documents uploaded onto Dspace must be carefully tagged with 'metadata' codes, such as subject keywords, which will help search engines to navigate the database.

DSpace should also spare MIT researchers much of the hassle of setting up and maintaining their own websites. The system supports a huge variety of file formats, allowing users to make anything from preprints to medical images available online. It also lets users attach other features to their documents, such as software that can manage access to restricted material.

But Smith warns that getting busy faculty members to provide high-quality metadata on their material is a big obstacle, and could deter many would-be participants. "You find that users are not beating down your door," says Eric Van de Velde, director of library information technology at the California Institute of Technology, who has developed a similar, smaller-scale institutional archive.

To get enough data onto the system to

make it useful, DSpace has grouped content around individual departments or labs that have well-organized publishing activities. The launch version of DSpace includes five of these 'early adopter' communities such as MIT's Department of Ocean Engineering and MIT Press. The groups have already placed about 750 documents on the database.

Smith says that a major demand from researchers is the storage of their raw data, such as medical images and other large data files. "A lot of these deserve to be managed and preserved as much as the publications that result from them," she says. She is also talking to publishers about the possibility of linking papers in their journals to supplementary information and data stored in the DSpace network of archives.

DSpace may also provide a more solid foundation for preprints than existing servers, but it can also support journals and new forms of information sharing, which could help libraries to cut down on journal subscription costs. In the long term, many academic centres feel that they need to retain control of more of what they produce, and DSpace may be a first step in that direction. ■

♦ www.dspace.org

Developing nations take initiative on greenhouse gases

Geoff Brumfiel, Washington

Some of the world's largest developing nations are already reducing their greenhouse-gas emissions — even though they have not yet been set targets. So says a report from the Washington-based Pew Center on Global Climate Change and the Battelle Memorial Institute, a private research organization based in Columbus, Ohio.

The report's authors say that this should be food for thought for the United States and Australia, which reject current attempts to limit their emissions in part because developing nations are exempt from such targets.

The study, which was released on 24 October, says that schemes implemented by Brazil, China, India, Mexico, South Africa and Turkey have together reduced the growth of their greenhouse-gas emissions by some 300 million tonnes per year over the past 30 years. The savings are the result of a wide range of programmes, from local renewable energy schemes to market reforms. In Brazil, for example, tax incentives are offered to families that buy smaller, more fuel-efficient cars, says William Chandler, an energy-policy expert who works for Battelle at the Pacific Northwest National



Solar power is just one way that developing countries have cut greenhouse-gas emissions.

Laboratory in Richland, Washington.

Such projects are voluntary under the first phase of the Kyoto Protocol, the international agreement on limiting greenhouse-gas emissions. Developing countries won exemption from emissions targets on the grounds that limits might hinder growth and that the developed world is historically responsible for the build-up of greenhouse gases. The United States and Australia cite this as a reason for rejecting the treaty (see *Nature* 390, 215–216; 1997), but Chandler says the report shows that some progress is being made without targets.

The report may also aid negotiations for the second phase of the protocol, which will cover emissions targets after 2012. Developing countries are likely to have to accept targets for this period, and they will need to have a quantitative understanding of their emissions to do so, says Tony La Vina, a senior fellow at the World Resources Institute, an environmental think-tank in Washington DC. "This study shows that it's possible," he notes.

But the report will probably not be enough to convince opponents of Kyoto that developing countries are doing their bit, says Mahendra Shah, a senior scientist at the International Institute for Applied Systems Analysis, an interdisciplinary think-tank in Laxenburg, Austria. "I think opponents will brush this report off because it is not comprehensive enough," he says.

In many developing countries, Shah points out, agriculture and forestry are the primary sources of greenhouse gases, but these sectors were not fully investigated in the report. Without a clear understanding of how these areas contribute to the picture, it will be difficult to convince the United States that developing countries really are working to reduce emissions, he says. ■

Haplotype mappers set their sights on genetic diseases

Washington A US- and Japanese-led global consortium this week launched the three-year, US\$100-million HapMap Project to chart genetic variation across the entire human genome. It is hoped that the venture will speed up the unravelling of genetic components of human health and disease (see *Nature* **412**, 105; 2001).

The HapMap will be constructed by sequencing DNA from blood samples taken from 200–400 Chinese, Nigerian, Japanese and American people to identify 'haplotype blocks', chunks of genetic markers that are sufficiently closely bunched on a chromosome to be inherited together. The project will build on efforts that have identified millions of single-nucleotide polymorphisms (SNPs) — sites where the genetic code differs between individuals by just one base.

Identifying the SNPs that tend to occur in each haplotype block yields more manageable clusters, making it easier to compare genetic data among groups. People with Alzheimer's disease, for example, might show different haplotype blocks from those without the condition, allowing scientists to

focus on these in their search for genes that are involved in the disease.

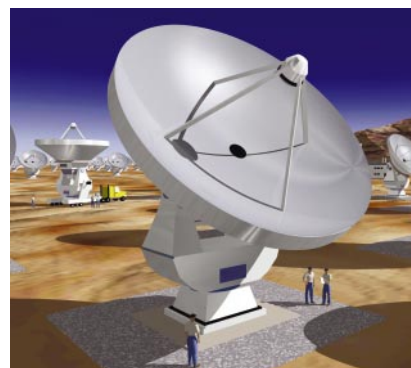
Kennewick Man's 9,200-year wait gets a little longer

San Diego Scientific examination of the 9,200-year-old skeleton known as Kennewick Man may be delayed by up to six years. The federal US District Court in Portland, Oregon, ruled on 22 October that four Indian tribes can appeal against a decision, made by a magistrate on 30 August, to allow archaeologists and anthropologists to examine the remains.

The August decision reversed plans by the Department of the Interior to give the skeleton — found in 1996 in the Columbia River in Washington state — to the tribes without scientific study (see *Nature* **419**, 5; 2002). The appeal could be rejected in about a year, but if it proceeds the process could take more than six years to complete.

Chile tunes in to host giant radio telescope

Washington The Chilean government last week approved the construction of the world's largest radio telescope, the Atacama Large Millimeter Array (ALMA), in the Atacama desert in the east of the country. ALMA will be made up of 64 dishes, each



Eyes on the sky: the Atacama array will study very distant galaxies and nearby gas and dust clouds.

12 metres in diameter, tuned to millimetre wavelengths, whose combined resolution is equivalent to a dish 14 kilometres across. It is expected to yield images ten times sharper than those from the Hubble Space Telescope.

On a plateau 5,000 metres above sea level, ALMA will have a view of the night sky undistorted by moisture. Researchers hope to use it to study very distant galaxies as well as clouds of gas and dust within our own Milky Way (see *Nature* **410**, 729; 2001).

The agreement, between Chile and the European and US project leaders — the European Southern Observatory, based in Garching, Germany, and Washington-based

Associated Universities Incorporated — is a milestone for the US\$660-million facility, says Paul Vanden Bout, interim director of the project and a researcher at the National Radio Astronomy Observatory in Charlottesville, Virginia. He hopes that the next phase, an environmental impact assessment and agreement on conditions for the involvement of Chilean astronomers, will be completed by early next year.

Europe fuels research into hydrogen-based energy

Brussels The European Union (EU) is to launch a bid to become the global leader in hydrogen-based energy technology.

A group of 19 senior representatives of the energy and transport sectors have been asked to come up with a 'foresight report' by 2003, which will outline an agenda for research on hydrogen and fuel cells, and strategies for using hydrogen in transport and industry. Hydrogen is considered the most viable alternative to fossil fuels. It produces only water when it is burned or used to generate electricity in fuel cells.

Romano Prodi, president of the European Commission, says that a substantial part of the 2.1 billion euros (US\$2 billion) earmarked for research on sustainable development, global change and ecosystems

under the EU's Sixth Framework Research Programme will need to be spent on hydrogen research in order to lead the field.

Agricultural science to tackle rural poverty

London The Washington-based World Bank is to explore the value of setting up a new international organization to assess ways in which agricultural science could help to reduce hunger and improve rural livelihoods in developing countries.

Options under consideration include the creation of an intergovernmental body along the lines of the Intergovernmental Panel on Climate Change (see *Nature* **412**, 112–114; 2001), or a non-governmental process akin to the Millennium Ecosystem Assessment (see *Nature* **417**, 112–113; 2002), or a hybrid of the two, says Robert Watson, chief scientist at the World Bank.

Watson is heading a global consultation that will consider the need for a new body, and its eventual scope and structure. The first meeting will be held in Dublin, Ireland, next week and will seek the views of scientists and consumers as well as governments and other organizations.

"The assessment must be demand-driven, not supply-driven. We must ask what is really needed," Watson says.



Putting on a good face: but drought has forced Zambian villagers to live on nuts and wild fruit.

Space physicist recreates the music of the spheres

Washington A musical composition, *Sun Rings*, featuring 'sounds' based on radio waves recorded by space probes, premiered in Iowa City last weekend. It uses 40 years of data converted into sound by Don Gurnett, a University of Iowa space physicist. Composer Terry Riley put the sounds to music, played by the San-Francisco-based Kronos Quartet.

Space is a noisy place at radio and other wavelengths. When the solar wind hits the ionosphere of Jupiter and other planets, for example, it generates 'choruses' that sound like birds chirping. The show will go on to tour in the United States and Europe.

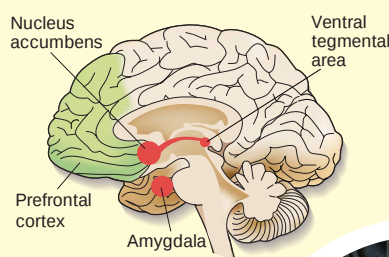
"When you are addicted, there is no euphoria when you shoot up," explains Christian. "You only want heroin. Food and sex are not interesting. You are capable of being aroused, but you have no desire." Can neuroscientists explain why addicts feel this way? Alison Abbott investigates.

Addicted

Christian, a heroin user from Munich, is proud to have been clean for two weeks. Like all reformed addicts, he has a story to tell. Most talk of how drugs are seductively rewarding at first, yet how the craving that follows has little to do with pleasure. But despite these common responses, there are big holes in our knowledge of addiction. Although researchers understand part of the neural circuitry involved, they know little, for example, about what changes in people's brains as they turn from recreational users into addicts.

To provide answers, neuroscientists are embarking on a new generation of experiments. Armed with sophisticated brain-imaging machines, researchers have launched a hunt for the neural signatures of behaviours linked to drug-taking, and are trying to correlate them with specific genes. To do so, they will have to work with much larger numbers of addicts — and their families — than previous studies have done. But if successful, they could help to fill the largest gaps in our understanding, such as why some people become addicted to drugs whereas others don't. They may even shed light on the biggest question of all: will neuroscientists ever be able to offer a cure for addiction?

The anatomy behind the initial drug high is already reasonably well understood, thanks to a series of ground-breaking experiments conducted by psychologist James Olds at McGill University in Montreal and the University of California, Los Angeles, during the 1950s. Olds inserted electrodes into the brains of rats, and allowed the rats to stimulate a particular brain area electrically by pressing a lever. The rodents liked doing it so much that they neglected to eat or drink¹.



Pleasure and pain: addictive drugs administer their effects to different parts of the brain.

The rats were exciting the mesolimbic dopamine system, an area now also known as the 'reward circuit'. This circuit consists of the ventral tegmental area, part of a region at the top of the spinal cord called the brain stem. Cells here send projections to higher emotional and cognitive areas, including the nucleus accumbens, a knot of neurons in the limbic, or emotional, brain (see diagram, above). Neurons throughout the system exchange signals by releasing a chemical called dopamine, which excites neighbouring neurons by binding to receptors on their cell membranes.

The reward circuit is part of the primitive motivational system in mammals, the normal function of which is to ensure that important behaviours such as eating and having sex are perceived as rewarding, thus increasing the likelihood that they will be repeated. Sweet tastes, for example, feed into the circuit through dedicated taste receptors

on the tongue, which send signals to the brain stem. A caress activates the system through sensory receptors in the skin.

But addictive drugs activate the circuit directly, and with dramatic effect. Cocaine blocks the protein that removes dopamine from the space between neurons, whereas amphetamines stimulate dopamine release.

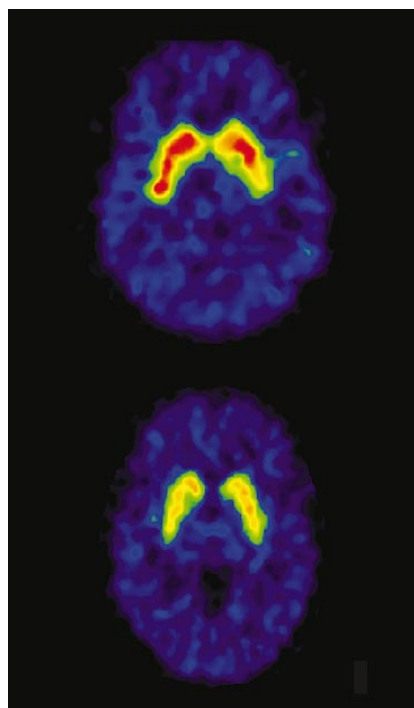
Heroin and nicotine activate the circuit by binding to other, non-dopamine

receptors on neurons, although how these receptors interact with the dopamine circuitry has not been worked out in detail. But whichever way they function, all drugs work by raising levels of dopamine in the nucleus accumbens to unnatural highs, and at unnatural speeds.

As work in animals has shown, drugs hit the mesolimbic dopamine system like a sledgehammer. "A natural rewarding activity such as sex or food will increase dopamine levels in the nucleus accumbens by 50–100%," says Roy Wise, a psychologist who pioneered research into reward circuitry and is now at the National Institute on Drug Abuse in Bethesda, Maryland. "A dose of cocaine or amphetamine can increase it by up to 1,000-fold."

Imaging studies show that drugs also activate the reward circuit in humans², but this is far from being the whole story of addiction. Addicts say that a stubborn craving for their drug develops, which is unrelated to its initial euphoric effects. "Pleasure is not what drives drug intake once addiction has kicked in, as any addict will tell you,"





Nora Volkow, seen at her PET scanner (right), has shown that a methamphetamine addict's brain (above, bottom) has smaller numbers of a specific dopamine receptor than normal.

agrees Nora Volkow, a neuroimaging pioneer at Brookhaven National Laboratory on Long Island, New York, who has worked with addicts for 20 years.

So what happens to the brains of drug users as this change takes place? Researchers agree that drugs must cause long-term changes, although the details are far from clear. Some evidence comes from cell-culture experiments and animal studies of several brain areas, including the reward circuit³. Both techniques have shown that prolonged administration of drugs alters gene expression, which leads to abnormal protein production. In addition, there is a fall in the density of the neurons' dendritic spines, the protrusions that act as sites for the junctions between cells.

Poor reception

The number of dopamine receptors in the reward circuit also changes. In 1997, Volkow described how she used positron-emission tomography (PET), which traces the movements of molecules tagged with a radioactive isotope, to study dopamine receptors in the reward circuits of methamphetamine and cocaine⁴ addicts. She found that the addicts had a lower level of one type of receptor, and that this reduction persisted for up to four months after they stopped taking the drug. "We haven't been able to continue a study for more than four months because addicts have usually relapsed by this time," she notes. Taking



cocaine may still elevate an addict's dopamine levels, says Volkow, but the reduced number of receptors prevents the drug from working its magic.

As well as tracking the changes that underlie addiction, researchers are probing the involvement of learning and memory. On the everyday level, we are used to links between memories and rewards. A whiff of cocoa can prompt a craving for chocolate, for example, and addictive drugs can forge stronger associations. Months or years after an addict is clean, craving can be reawakened by the sight of a needle or a visit to old haunts.

But only tentative hints exist about how these associations are learned. Cocaine and nicotine, for example, are known to alter the strength of connections between neurons — a phenomenon associated with learning — in the nucleus accumbens³. And several researchers have used imaging studies to show that drugs activate parts of the prefrontal cortex (PFC) — a complex structure at the front of the brain — that are involved in learning and laying down memories.

The involvement of this and other parts of the PFC has led to different theories about the development of addiction. Volkow is one of several researchers who have shown that the orbitofrontal cortex, part of the PFC, is consistently activated by drugs⁵. If a mild stimulant is used to induce craving in addicts, the extent of activation in this area also correlates with the intensity of craving. Volkow points out that the orbitofrontal cortex is dysfunctional in psychiatric disorders that are characterized by obsessive behaviours, and suggests that a similar dysfunction could underlie the obsession that addicts have with drugs. And Edythe London of the University

of California, Los Angeles, thinks that problems with the PFC's decision-making function are important in addiction, as addicts seem to lose their ability to weigh up advantages and disadvantages in their drug-seeking and drug-taking behaviours⁶.

But if such ideas are to be evaluated, new kinds of experiments will have to be devised. Physiological measures of addictive behaviour need to replace current methods, which rely on asking addicts about their feelings. And to understand why only some drug users become addicted, researchers will have to find a way of getting a handle on the genes involved, which means studying larger numbers of addicts and their families.

Cerebral snapshots

In the United States, new technology provided by the White House's Counterdrug Technology Assessment Center, which oversees the research programmes of the federal drug-control agencies, has kick-started such experiments. The centre is equipping the top US drug-addiction laboratories — about a dozen have so far been selected — with state-of-the-art imaging equipment, including PET and functional magnetic resonance imaging (fMRI) machines. PET machines generally provide the best spatial resolution, but can take up to two hours to complete a reading. Using fMRI, which tracks neural activity by monitoring blood flow in the brain, readings can be taken in minutes.

The machines underpin a new technique that could improve the way in which the mental states of addicts, and of those suffering from other psychiatric disorders, are measured. Over the past couple of years, neuroscientists have used imaging

techniques to search for endophenotypes, standard physiological responses that are observed in subjects carrying out particular mental tasks. The approach has already thrown light on the biology of depression.

Wayne Drevets of the National Institute of Mental Health in Bethesda has studied neural responses to sad faces, which are known to activate the amygdala, a brain area that is involved in processing negative emotions. Using PET scans, Drevets found that the response of the amygdala in depressed people continued unabated after several exposures to the sad face, whereas that of a control group wore off⁷. As the endophenotype — activation of the amygdala — was similar for all depressed subjects, Drevets hopes that it can be used as a potentially more accurate alternative to the standard list of psychiatric criteria that are currently used for diagnosis.

The hoped-for precision of endophenotypes could also make it much easier to identify genetic links to mental conditions. Proof of principle has been provided by Daniel Weinberger and Michael Egan, also of the National Institute of Mental Health, who have worked with an endophenotype for one of the cognitive deficits associated with schizophrenia — an abnormal activation observed in the prefrontal cortex of schizophrenics performing a memory test.

Genetic links

Last year, Weinberger and Egan showed that the presence of the endophenotype in schizophrenics and their relatives is linked to the gene that encodes an enzyme — catechol-O-methyltransferase, or COMT — that breaks down dopamine. The gene comes in two forms, one of which encodes a more active version of the enzyme. Weinberger and Egan found that subjects who displayed the endophenotype were more likely to have a higher proportion of the gene for the more active enzyme⁸. The effect was slight — equivalent to a 1.5% increase in risk of schizophrenia — but the association had not been picked up in studies that used less precise classical diagnostic criteria, despite the knowledge that schizophrenia is caused by dopamine dysfunction.

Neuroscientists now hope that the same approach could provide a clearer understanding of the genetics and neurobiology of addiction. Hans Breiter of Harvard Medical School, together with Harvard neurogeneticist Greg Gasic, plan to develop addiction-related endophenotypes for tasks associated with reward, for example. Breiter has received a fMRI machine that can produce a very powerful 7-Tesla magnetic field, which improves spatial resolution, from the counterdrug technology programme. “Having a 7-T fMRI machine — the most powerful machine available for human work — will be extremely valuable for this,” says Breiter. “The fine resolution will allow us to pinpoint areas of activa-



Greg Gasic (left), Larry Wald and Hans Breiter (at back) plan to seek out genetic clues to addiction.

tion in the brain much more accurately.”

Breiter and Gasic will begin by focusing on endophenotypes that should shed light on how activity in the reward circuit may be altered in addicts. They will use a series of tasks, which they believe could activate different parts of the reward circuitry⁹. These tasks will measure, for example, evaluation of social stimuli and response to monetary reward. In the latter task, for example, subjects will be scanned while watching a roulette-like wheel spinning and then stopping, and then learning whether they have won or lost a dollar sum. By comparing addicts and non-addicts, the researchers aim



Daniel Weinberger has found a genetic link to schizophrenia.

to pin down exactly how brain activity in the two groups differs in response to the various outcomes. “Asking drug addicts to report their feelings subjectively is too imprecise,” says Breiter. “Endophenotypes are going to be more helpful than subjective description of feelings, or even objective behavioural descriptions.”

In the initial stage of the project, a few hundred nicotine and cocaine addicts will be scanned to validate the endophenotypes. Next, up to 5,000 addicts and non-addicted members of their families will be studied to see whether the endophenotypes are heritable. Finally, a large, multi-centre study will be launched, based on these heritable endophenotypes, to identify genes that may predispose people to addiction.

The approach of the collaboration will be very much ‘top-down’. Instead of looking for candidate disease genes as in Weinberger’s ‘bottom-up’ approach, the consortium will look for single-nucleotide polymorphisms — changes of a single letter in the genetic codes of different versions of the same genes

— that are associated with the addiction endophenotypes. The approach will allow the researchers to search for links to the multiple genes that are expected to be involved in susceptibility to addiction — and also to spot unexpected links to genes that current theories do not flag up as likely to be important.

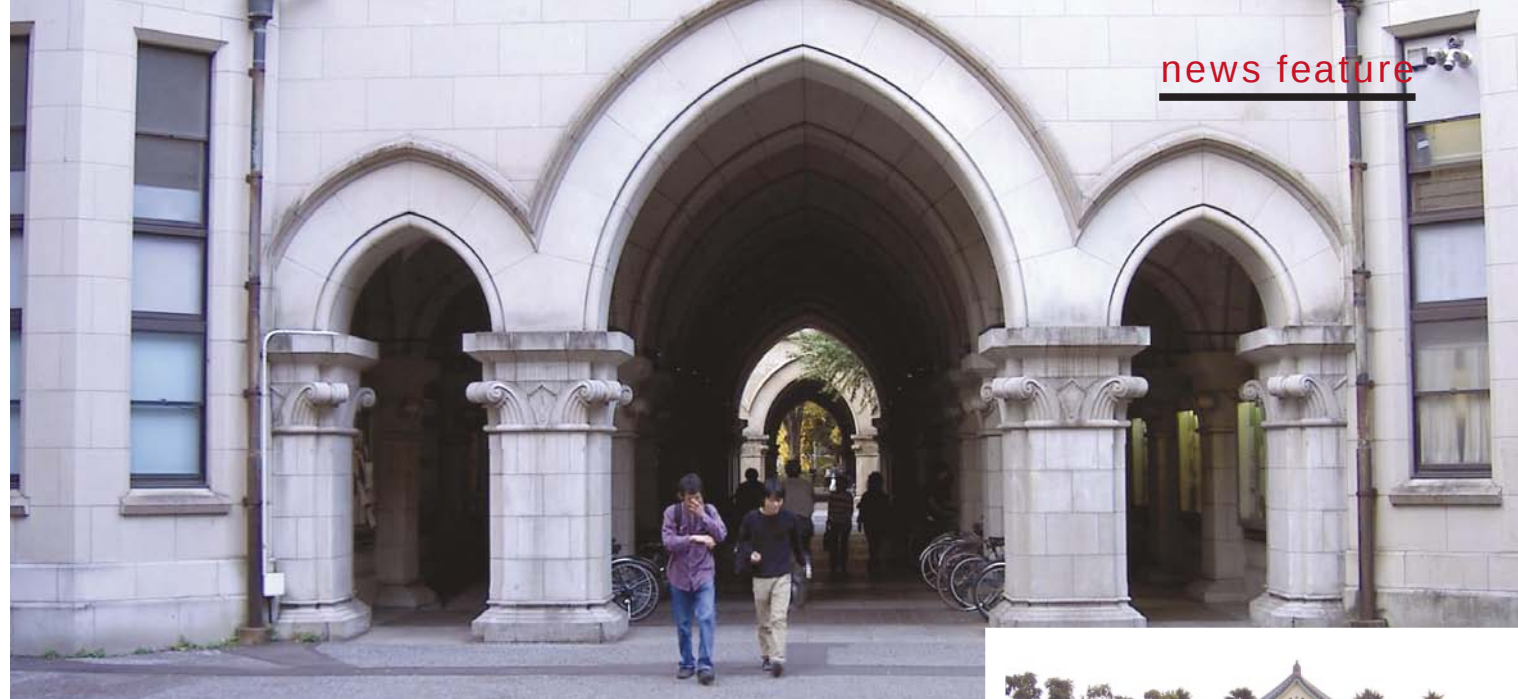
But it also means that much greater numbers of subjects are likely to be needed if trends are to be spotted amid the normal genetic variation of the population. Tens of thousands of people could eventually be required. “Our strongest challenge will be to scale up the pace at which we do high-throughput neuroimaging,” says Breiter. Most fMRI studies have been conducted with a few dozen, not a few thousand, subjects. This means that more computational infrastructure will be needed, as well as a massive step-up in computing power, which Breiter hopes that the counterdrug programme will also support.

The work of Breiter and his colleagues will take many years to deliver answers about susceptibility to addiction. But thanks to the increased availability of high-resolution brain-imaging machines, addiction researchers should, in the meantime, begin to clarify what they mean by terms such as ‘reward’ and ‘craving’. Whether a cure for addiction will follow is anyone’s guess, but a better knowledge of the brain mechanisms involved will certainly be an important step towards that goal. As the old adage goes — understand the biology and you’ll know where to target your medicine. ■

Alison Abbott is Nature’s senior European correspondent.

1. Olds, J. *Science* **122**, 878 (1955).
2. Spanagel, R. & Weiss, F. *Trends Neurosci.* **22**, 521–527 (1999).
3. Nestler, E. J. *Nature Rev. Neurosci.* **2**, 119–128 (2001).
4. Volkow, N. D. et al. *Synapse* **14**, 169–177 (1993).
5. Volkow, N. D. & Fowler, J. S. *Cereb. Cortex* **10**, 318–325 (2000).
6. London, E. D., Ernst, M., Grant, S., Bonson, K. & Weinstein, A. *Cereb. Cortex* **10**, 334–342 (2000).
7. Drevets, W. C. *Ann. NY Acad. Sci.* (in the press).
8. Egan, M. F. et al. *Proc. Natl Acad. Sci. USA* **98**, 6917–6922 (2001).
9. Breiter, H. C., Aharon, I., Kahneman, D., Dale, A. & Shizgal, P. *Neuron* **30**, 619–639 (2001).

White House Counterdrug Technology Assessment Center
 ♦ www.whitehousedrugpolicy.gov/ctac



D. CYRANOSKI

Independence days

Japan wants to reform its university system, in part to match the competitive and entrepreneurial spirit of US academia. That won't be easy, says David Cyranoski.



K. KAMOSHIDA/GETTY IMAGES (LEFT); KYODO

Standing before the upper house of Japan's parliament in May 2001, Prime Minister Junichiro Koizumi stunned officials from the education ministry by saying that he wanted the country's 99 national universities to be cut loose from state control. "We should privatize ones that can be privatized, and hand over ones that can be handed over to local governments," he asserted.

The universities had been widely criticized for not being competitive internationally. Koizumi's remarks were in response to a parliamentarian who noted that the annual survey on the competitiveness of 49 countries, compiled by the business school IMD International in Lausanne, Switzerland, ranked Japan last in terms of university education and entrepreneurship.

In the end, the universities will not be privatized, but Koizumi's comments have added impetus to reforms that should see Japan's universities change fundamentally over the next two years. The plan — named the Toyama plan after education minister Atsuko Toyama — has three pillars: giving universities administrative autonomy; making them compete to host research 'centres of excellence'; and consolidating the system through mergers.

There is general agreement that Japanese academia would benefit from increased competition. But is the system ready for such



All change: Junichiro Koizumi (left) and Atsuko Toyama want to transform Japan's universities.



an upheaval? Many researchers say that the plan will only succeed if it is backed by a solid system to evaluate research and teaching — which they fear is not yet in place.

What's more, some scientists are suspicious of the government's motives. They fear that the plan is an effort to save money and give academia a commercial focus. "The most important thing is to decide the purpose of the reform," says Nobel laureate Ryoji Noyori, a chemist at Nagoya University. "They should aim for energizing researchers and educators, not producing some financial result."

Currently, Japan's national universities

Could the domination of Japan's education system by the University of Tokyo (above) be enhanced by the proposed mergers of smaller institutes?

are agents of the state, and their staff are civil servants. But in 2004, they will become independent organizations. Rather than answering directly to the education ministry, each will report to two governing councils, one of which will take about half of its members from outside the university. The universities will still get government funding, but how much will depend on their success relative to other universities, and they will also have new freedom to get funding from private sources.

University presidents, formerly figureheads elected by faculty members, will be directly responsible for their university's success — like a US university president or a corporate chief executive. "Until now, everything has been done by consensus," says Masahiko Endo, president of Hirosaki University in Aomori, northern Japan. "Now presidents will be chosen for having a policy that can balance research and education and make the university succeed." Presidents will be selected by a committee formed from members of the two governing councils — although it is not yet clear whether new presidents will be appointed at all of the universities from day one.

Already, existing university presidents

have seized on the new 'centres of excellence' programme as a chance to preview their coming authority. More than 450 groups, each comprising about 20 professors, from 163 public and private universities applied to work on strategic, often interdisciplinary, research problems. Earlier this month, 113 teams at 50 universities were chosen to receive grants of between ¥100 million and ¥500 million (US\$800,000 to \$4 million) per year (see *Nature* **419**, 547; 2002) for five years.

The scheme was designed to encourage more competition between universities, while promoting interdisciplinary collaboration within each institution. "We broke the barriers between departments," says Akira Shimokohbe, vice-president of the Tokyo Institute of Technology.

In some cases, university presidents gave a nudge to reluctant faculty members. One molecular biologist at a university near Tokyo says that his group had little chance of being named as a centre of excellence, but spent the summer working on a proposal anyway. "It would have been better to wait until next year, but the president said we had to do it," he says.

Responsible care

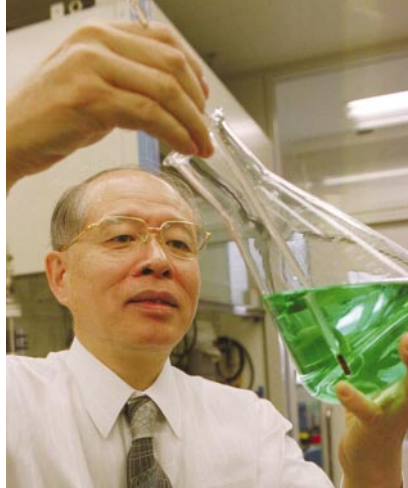
The vigour with which university presidents embraced the centres of excellence scheme suggests they are happy with increased competition. Most academics are also enthusiastic about reforms that allow them to carry over unspent portions of their equipment budget to subsequent years, or to hire staff without having to obey civil-service restrictions.

But for senior researchers, such freedom comes at a price. The loss of civil-servant status also applies to professors, who will be evaluated periodically to check that they still deserve their elevated position. If they are worried, most are not letting on. "We have already been proving ourselves to peer-reviewed journals and grant committees, so this won't be anything new," argues Mitsuhiro Yanagida, a molecular biologist at Kyoto University.

Officials at the education ministry hope that the new system will encourage universities to headhunt top researchers, who in turn will look for the best place to pursue their career. In the past, Japanese universities have been infamous for their 'inbreeding', with researchers staying at the same institution for their whole career. "Younger people can no longer expect to stay at one place for life," says Shimokohbe. Indeed, some researchers predict that up-and-coming scientists will soon start moving to universities to join a centre of excellence.

But these trends are causing some alarm at smaller, regional universities. They fear losing staff and being pressured into unequal mergers. The Toyama plan calls for a reduction in the number of universities, perhaps by as many as half, through amalgamations.

Already, some 12 mergers involving 24 universities are under way, with many more being considered. The mergers aim to share



Ryoji Noyori worries that the reforms are aimed at financial gain rather than academic excellence.

administrative costs and to improve the position of smaller universities in fields such as biomedical sciences by, for example, bringing medical schools together with universities that are strong in biological sciences but lack a hospital.

But critics fear that the merger mania could kill off pockets of strong research in regional universities. The reforms could also further concentrate resources in a system that is already dominated by the University of Tokyo and six other former imperial universities. "We need a system with many peaks," warns Yuichiro Anzai, president of Keio University, a private university in Tokyo.

Measure for measure

With a solid evaluation system, resources would naturally move to researchers doing the most innovative research — wherever they are working. But it is the readiness of Japanese academia to conduct thorough evaluations that is the main focus for concern over the Toyama plan.

Tsutomu Kimura, president of the National Institution for Academic Degrees, which has been developing a system to evaluate universities on how well they meet their own teaching and research goals, admits that the practice in Japan is immature. "Until now there have been 'soft evaluations' on a friendly basis," he says.

Many researchers say that funding bodies have fallen back on a system of rewarding well-connected researchers at major universities. "The problem is that the funding system is mostly top-down, and the people at the top are often retired researchers or bureaucrats who are out of touch with current research," says one biologist.

Some scientists argue that the best hope of breaking the 'old boy' network is to involve foreign researchers in evaluations. "It would only work if mostly foreigners were used," claims Yanagida. Tei-ichi Sato, director of the Japan Society for the Promotion of Science, one of the country's main research granting agencies, believes that applications for funding will increasingly be made in English to expand the scope for international peer review. But how

soon the system as a whole will embrace evaluation by foreign experts, and what will happen in the meantime, are still open questions.

Other scientists are worried that evaluation schemes will fail to recognize excellence in teaching, and base decisions on universities' research records alone. "The universities are not all the same," says immunologist Tadamitsu Kishimoto, president of Osaka University. "We may have to make a division between research and teaching."

Kishimoto and others also fear that evaluation will be connected too closely to government objectives to promote fields that may be crucial to Japan's future economic competitiveness. "It will be very dangerous if this competition is only geared towards fashionable science such as information technology, nanotechnology or proteomics," says Kishimoto.

The government certainly sees the universities as contributors to a strong economy. Last month, the education ministry decided to put a technology-transfer office — devoted to commercializing research discoveries — on each national university campus. So if researchers must tailor their projects to government priorities for economic growth, will the universities really become autonomous? "The reforms are likely to make universities, especially middle-ranking ones, more dependent than independent," claims Yanagida.

Almost everyone agrees that Japan's university system has much to learn from US academia. But although the government's reforms are designed, in part, to recreate elements of that system, critics point to important differences that will make it hard for Japanese universities to become genuinely autonomous and entrepreneurial. Most importantly, the United States enjoys a tradition of private donations from individuals that is not at all developed in Japan.

In April, for instance, Harvard University received a donation in the will of chemist Herchel Smith, one of the pioneers of the contraceptive pill, which brought his contributions to the university to about US\$100 million — four times what Osaka University receives in an average year. "Such donations from individuals are unheard of in Japan," laments Kishimoto.

David Cyranoski is *Nature's* Asian Pacific correspondent.

D. CYRANOSKI



A list of published papers is no measure of value

The present system rewards quantity, not quality — but hasty changes could be as bad.

Sir — The choice of performance indicators sends a powerful message to those being evaluated, and when those measures are linked to the distribution of research funds, academics are quick to respond. Our analysis of Australian university publications shows clearly how the sector has reacted to funding formulae that reward quantity rather than quality.

A large part of the government funds that support the research activities of Australian universities is allocated on the basis of formulae that comprise three elements: research income, postgraduate students and publications¹. Data on the third element have been collected annually since 1993. When this element was incorporated into the funding formulae in 1995, universities and researchers were quick to calculate the 'value' of a publication.

Between 1995 and 2000, this figure varied from A\$761 (US\$415) to A\$1,089, influenced by the publication types included and the total funds allocated. After a review of higher-education research in 1999, the amount to be distributed via formulae increased significantly, to more than half of the funding specifically targeted to research and research training through the Education, Science and Training portfolio. As a result, a published paper is now 'worth' more than A\$3,000 to a university. The value to individuals or departments can be appreciably higher.² Some universities distribute these funds internally using the same formula, but giving more weight to publications — up to three times the sector value.

We have quantified the apparent effect of this policy. We have distributed all journals from the Institute for Scientific Information's (ISI) *Science Citation Index* (SCI) into quartiles, using journal impacts calculated on the basis of five-year citation means, and have tracked the presence of Australian universities in these four journal sets over time (Figure 1).

The response of the academic community was predictable and clear. Until the 1989–93 period, there had been virtually no movement in the institutions' presence in the SCI journal sets. Subsequently, university output jumped noticeably, even though funds remain extremely tight and academic staff numbers stable. The most striking feature of that increase has been its lack of uniformity. The sector's share of publications in the top two quartiles rose by around 20%, but in the third quartile

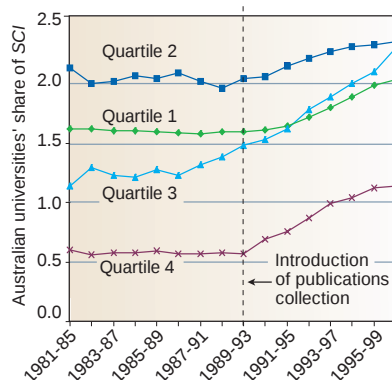


Figure 1 Australian universities' share of publications in the SCI, by journal impact quartile: five-year windows, 1981–1985 to 1996–2000.

its share increased by over 50%, and it doubled in the bottom quartile.

With no differentiation between the quality or impact of the publications, there is little incentive to strive for placement in a prestigious journal. Whether a publication is a groundbreaking piece in *Nature*, or a pedestrian piece in a low-impact journal, the rewards are identical. And with the recent trebling in the 'value' of a publication unit, the

pressure to focus on this will not diminish.

Concerns that this component of the funding formula was not measuring the characteristic that it was designed for — quality — were raised soon after its introduction. However, not all universities were keen to see it removed or replaced. For smaller institutions, this particular element was more rewarding, and easier to improve, than the others.

These concerns are now re-surfacing in the context of the latest review of the Australian higher education system³. A number of submissions to recent ministerial discussion papers have suggested the removal or modification of the publications component. The difficulty is that suggested alternatives are as problematic as the one they seek to replace. It is to be hoped that time will be taken to analyse the likely effects of any alternative measures before they are introduced.

Linda Butler

Research Evaluation and Policy Project, Research School of Social Sciences, The Australian National University, ACT 0200, Australia

1. <http://www.detya.gov.au/highered/research/index.htm>
2. http://www.avcc.edu.au/policies_activities/resource_analysis/key_stats/kstats.htm
3. <http://www.detya.gov.au/crossroads>

Realistic attitude takes postdocs a long way

Sir — Compared with other countries in Asia, Japan is often described in discouraging terms with respect to foreign researchers working there (see, for example, *Naturejobs* 4–5, 8 August 2002). I have just completed four years' postdoctoral work in Japan, and agree that it is more difficult for Japan than for the United States or Europe to attract young foreign researchers. This is a great pity, as both the country and its science have a lot to offer postdoctoral fellows.

One of the greatest obstacles is the image that Western scientists have of Japan. At a recent conference, numerous graduate students and postdocs asked me questions about my experience, almost always starting out with whether it is difficult to live there. No, it is not difficult — not least because the financial support provided by institutions such as the Japan Society for the Promotion of Science (JSPS) is very generous.

Two critical elements for a successful stay are the attitudes of the researcher and

of the laboratory head. Foreign researchers should develop at least basic Japanese language skills and not expect the lab to burst into English for every detail. Although good speakers may come to the institution, Japan is geographically distant from countries in which most scientific meetings are held, so visiting postdocs should make the most of the grants offered by the JSPS for travel to conferences.

The wonderful time I experienced in Japan was largely due to the supportive nature of the lab I worked in. Frequently, however, one hears stories where this is not the case. If the *sensei* (lab head) is unenthusiastic or is prejudiced against foreign researchers, conflicts arise. In my view, it is crucial for a researcher to meet his or her prospective lab head before deciding to move.

As PhD graduates rarely have the funds to travel at the end of their studies, a programme (perhaps funded by JSPS) to allow a visit to a prospective lab, with no commitment on either side, would be useful in establishing regular successful working relationships between foreign and Japanese researchers. Preparedness and

flexibility are the key to finding success and happiness while working in Japan.

Stuart Fraser

Laboratory of Molecular Mouse Genetics,
Institute for Toxicology, Johannes Gutenberg-
University Mainz, Obere Zahlbacher Strasse 67,
Mainz 55131, Germany

Science, conservation and fox-hunting

Sir — Much evidence on the issue of fox-hunting with hounds is either speculative, being based on questionnaire surveys, or contradictory, particularly where funds are provided by special-interest groups. The recent study done at Bristol University (P. J. Baker, S. Harris & C. J. Webbon, *Nature* **419**, 34; 2002) is noteworthy for attempting an experimental approach.

Baker *et al.* found that the temporary cessation of fox-hunting in Britain during the foot-and-mouth disease outbreak of 2001 had no impact on fox population density, and concluded that a permanent ban on hunting is unlikely to result in a dramatic increase in fox numbers. However, motor vehicles are the greatest killer of foxes in Britain, accounting for some 25% of deaths. Hunting with hounds accounts for only 6.3% of the 400,000 foxes killed annually. More than five times as many are killed by shooting and snaring as by hunting with hounds in lowland hunting areas (L. Burns, V. Edwards, J. Marsh, L. Soulsby & M. Winter. *Report of the Committee of Inquiry into Hunting with Dogs in England and Wales*, Stationery Office, London; 2000; see www.hunting-inquiry.gov.uk). Fox-hunting is an ineffective method of population control.

Instead, these data suggest that fox-hunting harvests a sustainable off-take, which might represent a traditional form of community-based conservation. Such projects improve local tolerance towards wildlife and maintain biodiversity without statutory regulation and recurrent public funding. The British government has supported many such projects in developing countries, and is committed to doing the same in Britain as a signatory to the Convention on Biological Diversity.

The defence of fox-hunting on conservation grounds relies on two main predictions in the event of a ban: first, that voluntary maintenance of biodiversity-rich fox habitats such as woodlands and hedgerows by landowners involved in hunting would decline; second, that landowners' tolerance of foxes would decline, increasing their persecution by other potentially less humane methods and so reducing fox numbers. Landowners may have the potential to reduce fox

densities by shooting and snaring (M. Heydon & J. Reynolds, *J. Zool.* **251**, 265; 2000), but using these results to predict changes after any ban remains problematic.

The best way to test these predictions would be to build on the opportunistic approach attempted by Baker *et al.* by imposing a temporary, medium-term ban in randomly chosen areas and conducting independently funded research into its effects on a range of factors. This adaptive management approach would satisfy Lord Burns's recent recommendation not to rush a decision on whether to ban hunting. Although this approach has its pitfalls, we believe that, with careful planning, it would provide a firmer scientific basis for legislation than existing evidence.

**N. Leader-Williams, T. E. E. Oldfield,
R. J. Smith & M. J. Walpole**

Durrell Institute of Conservation and Ecology,
University of Kent at Canterbury,
Canterbury CT2 7NS, Kent, UK

Culture gap: in biology, what works, continues

Sir — Despite the arguments reported in your News Feature "Bridging the culture gap" (*Nature* **419**, 244–245; 2002), biologists already have a simple unifying rule, without the help of physicists. It is 'what works, continues' — usually stated in terms of the survival and reproduction of the fittest.

In answer to a posed question: phosphate is used to activate and deactivate proteins, as are methyl and ethyl groups and various saccharides, for the same reason that I currently use green and orange highlighters. At some time in the past they were there and functioned, and were incorporated into the system. Applied maths has its place in biology, especially where simple rules apply, in detecting signal in noise and defining practical limits.

Mendel was the first and most influential in this regard. His work was so clever, or so arcane, that it took 35 years to work out what he had discovered, and another 50 years for molecular genetics to explain the mechanism that causes dominance. Typically, Mendel's laws underestimate reality. The effects of most alleles on most characters are quantitative, polygenic and multi-factorial, rather than qualitative — tall versus short.

Compare an organism to an automated factory. Physics can explain all of the functions from electrons in transistors to computers in robots to metal-forming stresses and welding, but it has trouble with company balance sheets and share prices.

In business, the overriding factor is market share; in biology, habitat occupation.

Hugh Fletcher

School of Biology and Biochemistry,
Queen's University of Belfast, 97 Lisburn Road,
Belfast BT9 7BL, UK

Culture gap: physics still seeks its unifying theory

Sir — I was somewhat bemused when reading your News Feature (*Nature* **419**, 244–246; 2002). The view that "biology today is where physics was at the beginning of the twentieth century" misses a critical difference between the two disciplines. Biology has a grand unifying theory: it was published in 1859 by Charles Darwin as *On the Origin of Species by Means of Natural Selection*. The same cannot be said of physics, which continues to search for its theory of everything.

D. J. Hosken

Zoology Museum, University of Zurich,
Winterthurerstr 190, 8057 Zurich, Switzerland

Patents limit medical potential of sequencing

Sir — In your interesting Nature Science Update¹ "24-hour genome dawn", you report on the prospect of a personal sequence in minutes, for less than \$1,000. Patents will present at least two major problems to the timely adoption of these technologies^{2,3}.

First, some US companies will not license 'their' genes for testing by others, so any diagnostic chip would have to skip the patented gene estate of Myriad Genetics and similar outfits. Second, for those willing to license their genes non-exclusively for inclusion in diagnostic gene chips and similar tools, the stacked royalties payable on all the patented genes will make the tests prohibitively expensive.

Technological advances will benefit patients only if owners of diagnostic gene patents permit the technologies to be used and are reasonable in their demands for royalties, such as by limiting their expectations to a small fraction, say 1–3%, of the marginal cost allocable to their genes⁴.

Jon F. Merz

Center for Bioethics, University of Pennsylvania,
3401 Market Street, Suite 320, Philadelphia,
Pennsylvania 19104-3308, USA

1. <http://www.nature.com/nsu/020923/020923-2.html>;
24 September 2002.

2. Schissel, A., Merz, J. F. & Cho, M. K. *Nature* **402**, 118 (1999).

3. Merz, J. F., Kriss, A. G., Leonard, D. G. B. & Cho, M. K. *Nature* **415**, 577–579 (2002).

4. Merz, J. F. *Clin. Chem.* **45**, 324–330 (1999).

Taking the subatomic Grand Tour

The secrets of the world of elementary particles revealed.

The Particle Odyssey: A Journey to the Heart of Matter

by Frank Close, Michael Marten &

Christine Sutton

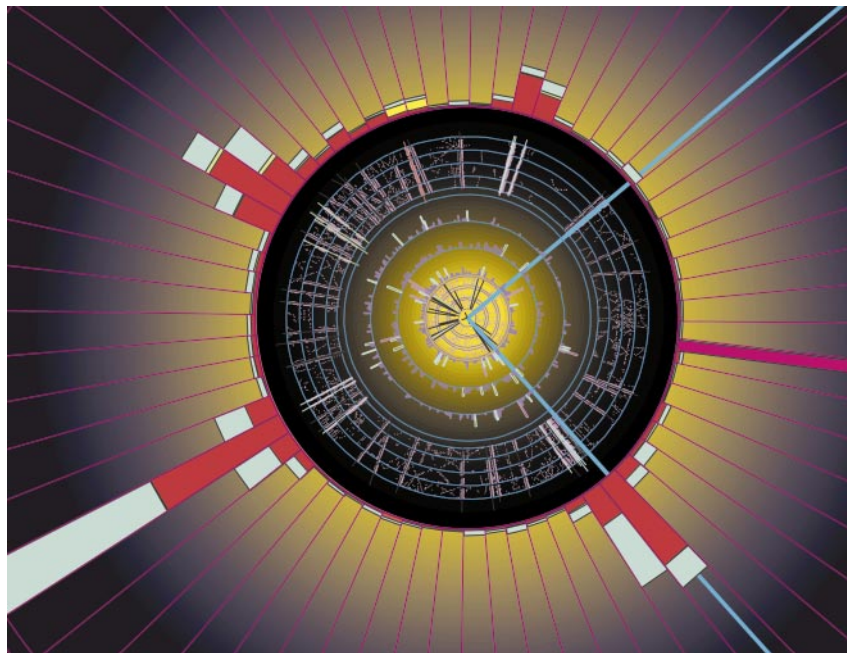
Oxford University Press: 2002. 246 pp.

£29.95, \$45.

Ken Peach

The development of the standard model of particles and their interactions is one of the major scientific achievements of the twentieth century. The electron was the first 'elementary particle' to be discovered, by J. J. Thomson (working largely on his own) in 1897, and the top quark was the last of the standard model's 'fundamental fermions' to be discovered, by two large teams at the Tevatron near Chicago in 1995. In between, special and general relativity, quantum mechanics and a deep understanding of symmetry (explicit and hidden, global and local, absolute and spontaneously broken) have transformed our understanding of the dynamics of the Universe. Within this framework, a remarkably complete description of the way in which the Universe works has been developed, and all but one of its ingredients experimentally verified — only the elusive Higgs particle remains undetected.

The Particle Odyssey describes a century of discovery and creativity through a series of stunning images and photographs. The text explains clearly and non-mathematically the significance of these discoveries in the development of the standard model. Each



Problem particle: the top quark, here decaying into muons (blue tracks), wasn't discovered until 1995.

ingredient (quarks, leptons, gauge bosons and other whimsies associated with high-energy physics) is described and illustrated, along with photographs and biographical details of key figures in the story.

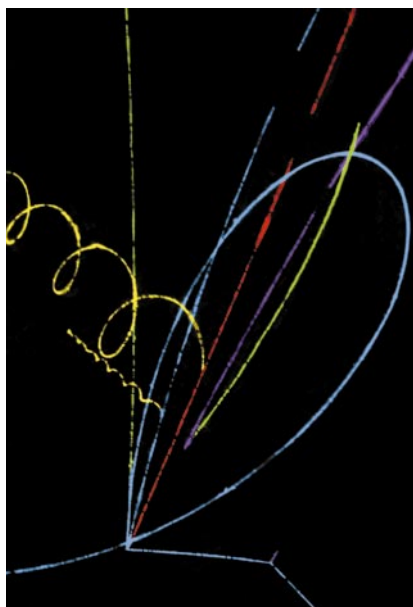
The book's 12 chapters cover broad themes, loosely following the historical development of the subject. Two chapters cover the technology — accelerators and detectors — and a third looks at the applications of these in other branches of science and society. The style is conversational, with the imagery sometimes as vivid as the photographs. Although there are some crucial or illustrative histograms and scatter plots (the 'picture' that the physicist really wants to see), the emphasis is on people, events and equipment. The book is written partly as history and partly as detective story, identifying 'who' as well as 'what', 'where', 'when' and 'why'. It will be a useful source for future historians of the science of the twentieth century, limited only by the lack of references to the original papers.

The present book builds upon *The Particle Explosion*, published 15 years ago by the same authors. Although its chapters share headings with the earlier book, it is more than 'Particle Explosion II — the Sequel'. It is easy to forget just how much progress there has been in high-energy physics over the past 15 years. The main ingredients of the standard model had been discovered by then, but there was still scope for alternative explana-

tions. Most of the loopholes have since been closed: the top quark has now been discovered, the electroweak sector has survived precision scrutiny at the Large Electron-Positron Collider at CERN, quantum chromodynamics has been explored in detail by the HERA experiments at DESY, both solar and atmospheric neutrinos have been shown to oscillate, the tau neutrino has been observed, hints of the quark-gluon plasma have been seen, and matter-antimatter asymmetries have been measured in B-mesons at the KEK-B and PEP-II B-factories. The interplay between high-energy physics and cosmology has been enhanced. All this is reflected in the new book, with about one-third of the text and more than 250 new images illustrating this progress. If *Explosion* conveyed the image of creative, if chaotic, energy, *Odyssey* gives the impression of a journey surveyed from a vantage point, from which one can see the route traversed so far and contemplate the road that lies ahead.

In short, *The Particle Odyssey* is a beautifully illustrated and eminently readable introduction to high-energy physics. It provides an excellent answer, for both the high-energy physicist and the general reader, to the party question that high-energy physicists sometimes find difficult to answer — what is it exactly that you do? ■

Ken Peach is in the Department of Particle Physics, CLRC Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK.



Gone in a flash: the charmed sigma particle is inferred from tracks left by other particles.

Reinventing the chemical industry

Handbook of Green Chemistry and Technology

edited by James Clark & Duncan

Macquarrie

Blackwell Publishing: 2002. 540pp. £149

Martyn Poliakoff & Peter Licence

Across the world, chemical manufacture is unsustainable. Its petroleum-based feedstocks are dwindling and the problems and cost of waste disposal are increasing rapidly. Environmental legislation is becoming ever more restrictive, and the industry has a poor public image. But modern society is almost totally reliant on the products of the chemical industry, and rising living standards in developing countries are increasing demand. We cannot just stop making chemicals. How then is chemical manufacturing to be transformed into a sustainable, non-polluting activity before the price of oil becomes prohibitive?

'Green chemistry' (see *Nature* **413**, 257; 2001) is a relatively new approach to resolving this impasse by removing the hazards of chemical usage through the design of safer manufacturing processes and safer chemicals. It aims to achieve sustainability by increasing the use of biologically derived, renewable feedstocks and by minimizing waste. Such broad aims require a correspondingly wide range of skills; consequently, green chemistry involves toxicology and life-cycle assessment as well as chemistry and process engineering. Reaching these goals also demands unusually close collaboration between academia and industry.

The *Handbook of Green Chemistry and Technology* is a multi-authored book that aims to illustrate the state of the art across

a wide range of green chemistry. It is not a revolutionary manifesto but rather a 'practical and rigorous' collection of reviews with references up to the end of 2000. The 23 chapters are the work of 28 authors from eight different countries but, disappointingly, only four of them are from industry. The coverage is broad, with chapters on most key aspects of green chemistry, including catalysis, solvents, reagents, life-cycle assessment and unconventional technologies such as ultrasound or microwaves for promoting reactions.

The chapter on 'Hydrogen Peroxide in Waste Minimization' quotes an industrial chemist who maintained that "almost all the chemistry really used was at least 50 years old". This is a key problem of green chemistry. Often, what is new is the way that reactions are used, rather than the reactions themselves. The most enlightening chapters are those such as 'Waste Minimization in Pharmaceutical Process Development', in which particular examples explain the logic and science behind cleaner chemistry. Other chapters fail to highlight the 'green' aspects of their topic and are little more than convenient summaries of catalysis or technology. The chapter on 'Green Chemistry in Practice' is the only one to tackle the problems of non-petroleum feedstocks in any detail. Processing such feedstocks is an area in which biocatalysis is likely to have a major role, so it is disappointing to find biocatalysis discussed in only 21 pages, hardly more than are devoted in another chapter to the solid catalysts from a single company.

There is a useful introduction to electrochemistry, albeit with an unnecessary overlap with the chapter on fuel cells, and a succinct description of the use of high-pressure (supercritical) carbon dioxide as an environmentally less harmful replacement for conventional solvents. Ionic liquids — salts that melt below room temperature — are also

being widely promoted as alternative solvents. Although these liquids are mentioned briefly in the context of phase-transfer catalysis, they warrant a chapter of their own. The book has an intriguing, abstract photograph on the cover, but readers have to wait until page 367 to discover that it is a "view of sheared thin films in a spinning disc reactor". This reactor is briefly covered in a chapter on process intensification, which involves making chemicals in small, high-throughput equipment — another subject that deserved more space. The final chapter, on the extraction of green products, ends with the intriguing statement that the extraction of green (unroasted) coffee beans with superheated water "produced a brown liquid with the aroma of coffee".

Overall, the book is a useful resource for those working in the area; it is not a guide for the novice green chemist. But it does show that there is more to green chemistry than merely using hydrogen peroxide to bleach your waste water so that people do not notice when you pour it into a river. ■

Martyn Poliakoff and Peter Licence are at the School of Chemistry, University of Nottingham, Nottingham NG7 2RD, UK.

Setting the record straight

Rossiiskaya Nauchnaya Emigratsiya: Dvadsat' Portretov (The Russian Scientific Emigration: Twenty Portraits)

edited by G. M. Bongard-Levin & V. E. Zakharov

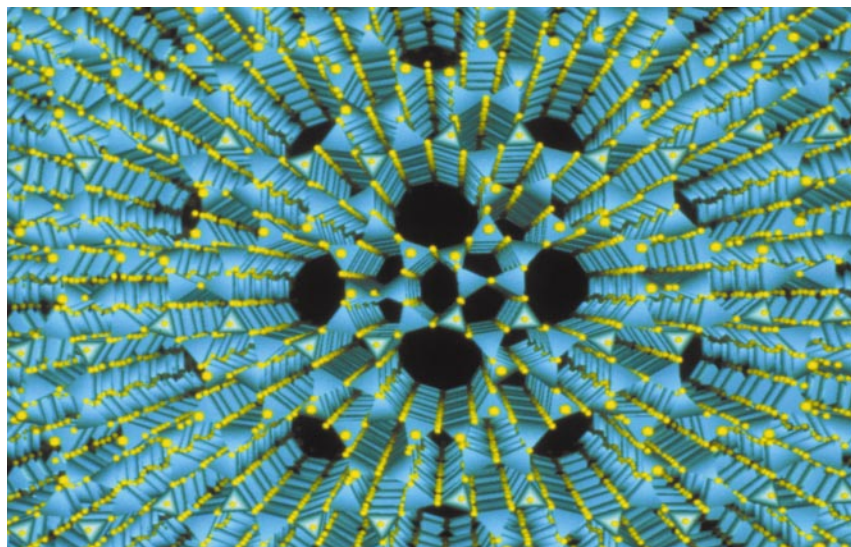
Editorial URSS, Moscow: 2001. 386pp.

18 euros

Valery N. Soyfer

In 1564, the Russian prince Andrei Mikhailovich Kurbsky quit the service of Tsar Ivan the Terrible (Ivan IV) and emigrated to Poland. An educated courtier, a military commander and later a historian and literary figure, Kurbsky was highly regarded by the Polish king Sigismund II Augustus. From 1564 to 1579, Kurbsky and Ivan the Terrible corresponded: Kurbsky accused the tsar of cruelty and tried to explain to him why he preferred voluntary exile over ordeals and possible death, while the tsar accused him of betraying the motherland and of anti-Russian attitudes.

Such animosity endured in Russia towards those who abandoned their country, reaching a peak during the Soviet period. After the 1917 revolution, nearly 2 million educated and talented Russians emigrated to the West. The names of many Russian émigrés have entered the history of the twentieth century, including Vladimir Zworykin,



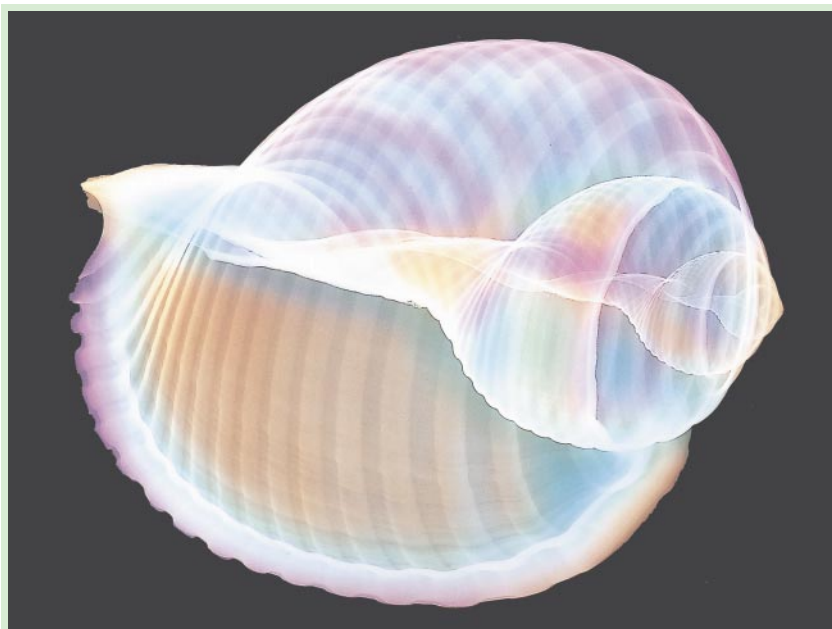
Green chemistry: zeolite minerals provide an environmentally friendly route to refining petroleum.

the inventor of television; Igor Sikorsky, the creator of the helicopter; and the composers Igor Stravinsky, Sergei Rachmaninov and Nikolai Metner. But the communist rulers did not recognize that these people were bringing glory not only to the countries that gave them shelter, but also to their motherland. The émigrés were slandered and declared to be traitors and enemies.

When, in the 1920s and 1930s, several leading scientists did not return to the USSR after travelling abroad, they began to be persecuted. The communists applied a harsh strategy to the most eminent scientists: elected members of the Russian (or Soviet) Academy of Sciences who emigrated were excluded from the list of members. This happened to the historian Mikhail Rostovtsev in 1927, and to the chemists Alexei E. Chichibabin and Vladimir N. Ipatiev later. Their names and those of the great majority of Russian émigrés were struck from the memory of Russians for nearly half a century.

However, a remarkable book has now been published in Russia on the lives and affairs of 20 of the most important Russian scientific émigrés who achieved worldwide renown not in their homeland, but in exile. Most of the protagonists of this book shared similar fates. They began in their new environments with difficulty, and none expected a warm reception. But possessing enormous creative potential and unusual strength of character, they made their way to the peak of the scientific Olympus in the West.

A number of Russian scientists went to America, among them the astronomer Otto Lyudvigovich Struve. Struve graduated from university in Russia in 1914. He fought in the White Army against the Bolsheviks but was seriously wounded and escaped to the West. His uncle, Paul Gutnik of the Babelsberg Observatory in Berlin, wrote to the director of the Yerkes Observatory at the University of Chicago, encouraging him to take his nephew



The eye of the beholder

Much of the world around us cannot be seen by the naked eye — we need techniques such as scanning electron microscopy or satellite imaging. For instance, the fine detail shown here of the coiled shell of the tun (*Tonna galea*), a marine carnivore, was obtained using false-colour X-ray

techniques. In *Heaven & Earth: Unseen by the Naked Eye* (Phaidon, £29.95, \$49.95, 49.95 euros), David Malin has compiled a stunning collection of images, which without such techniques would be lost to us, of objects ranging in size from a tiny gold atom to galaxies billions of light years away.

on. Once there, Struve began to use spectral methods in astronomy, pioneering work in several directions, especially in stellar spectroscopy (using absorption lines). He also discovered interstellar calcium and hydrogen. He became director of the Yerkes Observatory before heading the new Leuschner Observatory in Berkeley, California.

The physicist Georgiy Antonovich Gamow became a well known and colourful figure in world science. Gamow contributed

to three fundamental scientific fields of the twentieth century: physics (he discovered the quantum nature of α -decay), cosmology (he created the theory of a hot Universe, and on the basis of that he predicted the existence of the Big Bang) and genetics (he guessed the essential nature of the universal genetic code). Gamow and his wife were allowed to leave Russia for two weeks in 1933 to attend a Solvay Congress on physics, but never returned. At that time, failure to return to the USSR was regarded by the Soviet authorities as espionage, betrayal of the motherland, and conspiracy with the goal of seizing power. The death penalty was prescribed for such actions (this statute was removed in Russia only in 1995). Thus, Gamow would often tell acquaintances in America: "I am living under a sentence of death."

The physical chemist Georgiy Kistya-kovsky was a professor at Harvard University and also worked on the Manhattan Project to create the American atomic bomb. He became an important figure in the White House, advising President Eisenhower on issues of national security.

Other émigrés settled in Europe. Sergei N. Vinogradsky, considered to be the founder of modern soil microbiology, came to France in 1923. He organized a laboratory in the Pasteur Institute near Paris but stayed in contact with Soviet scientists and remained an honorary member of the Soviet Academy



Russian reunion: scientists gathered in Berlin for Russian Science Week in 1927 included the chemists Vladimir N. Ipatiev (standing, eighth from right) and Alexei E. Chichibabin (standing, fourth from right).

of Sciences. The author of the essay on Vinogradsky, G. A. Zavarzin, claims that he "continued to take an interest in scientific life in Soviet Russia", which meant maintaining friendly relations with the bosses of Soviet science, in contrast to other Russian émigrés. But this singling out of Vinogradsky is perhaps unjustified. As the other essays in this book show, most émigré scientists remained Russian patriots and helped many of their friends who remained in the USSR. They simply could not reconcile themselves with the morals and behaviour of the communists. They condemned the repressions and terror in their motherland, and tried to avoid the retribution of the Bolshevik hangmen.

Other Russian émigrés featured include Vladimir A. Kostitsyn, an astrophysicist who

also worked in mathematics and ecology, the mathematician Dmitry P. Ryabushinsky and Chichibabin, who all emigrated to France. Also profiled here are geneticist and evolutionist Theodosius G. Dobzhansky, mathematician Yakov (Jacob) Tamarkin, engineer Stepan P. Timoshenko and Ipatiev.

The final section is devoted to biographies of eminent specialists in the fields of economics and the humanities: the Nobel laureate Vasily Leontiev, the historians Rostovtsev, Georgiy Vernadsky and Pavel Milyukov, and the philologist Roman Yakobson, a Harvard University professor who had great influence in many fields of the humanities.

The book is full of detailed descriptions of the degrees, prizes and awards bestowed on the subjects of the essays, but unfortunately

there are no name or subject indexes and no bibliography. Only some of the authors indicate the sources from which they obtained their materials. These materials would be useful not only to researchers of Russian history, but also to show more fully how the history of Russia was viewed in the West.

Nevertheless, this book is the first work in which the lives and activities of representatives of those who achieved significant success during their lives outside Russia are discussed without any political subtext. Although it is primarily addressed to a Russian audience, it would be good to see it translated into European languages. ■

Valery N. Soyfer is in the Laboratory of Molecular Genetics, George Mason University, Fairfax, Virginia 22030, USA.

Science in culture

Shingu's showers

Technology and nature at Fiat's Lingotto factory in Turin.

Martin Kemp

On the face of it, there seems to be something incompatible between high-tech components in gleaming metal, obviously precision-engineered, and an insistence upon a total harmony with nature in all its living complexity. But any incompatibility is triumphantly dissolved by the kinetic sculpture of Susumu Shingu, the Japanese constructor of devices driven by natural forces.

His large-scale constructions grace public spaces throughout the world, including Queens Criminal Court in New York and the New England Aquarium in Boston. His travelling installations of sculptures, *Windcircus*, which toured urban sites in the late 1980s, and *Wind Caravan*, which was encamped in stunning natural locations in Japan, New Zealand, Finland, Morocco, Mongolia and Brazil during 2000 and 2001, are at once forceful and light, strong and elegant, large and delicate — evoking the power and sensitivity inherent in natural systems.

There could be no more telling context for him than Fiat's classic Lingotto factory in Turin, Italy, one of the paradigmatic structures of the modern era. Built to the master-plan of engineer Giacomo Mattè Trucco, starting in 1916, it is a great hollow rectangle 500 metres long. High on its roof is an automobile test-track with parabolic banking. Having reluctantly abandoned it as a manufacturing base in 1982, Fiat has progressively brought new life into the great space. Now, with the recent opening of the Giovanni and Marella Agnelli Art Gallery in a 'pod' extruded from the roof, the rehabilitation is complete. Renzo Piano — who came to world prominence when he and Richard Rogers won the design competition for the Pompidou Centre in Paris — is the architectural mastermind, and it was at his instigation that Shingu was invited to provide a water sculpture to stand in front of one of the building's long facades.

The 11.2-metre central column of Shingu's *Locus of Rain* consists of a hollow tube from which a water spray spouts high into the air. Around this revolve a series of arms, trumpet-shaped vessels and counterweights. Each of the two units on either side of the axis revolve in vertical planes around three pivots, the first on their shared axle, the second at the point where the lateral arms intersect with the poles that support the 'trumpets', and the third in the necks of trumpets themselves.

The ensemble is driven by water that falls into the trumpets, affecting not only their own rotation, but those of the poles that support them and the primary arms. Each mouth drinks from the spray when it swings towards the vertical, and turns gracefully aside as it becomes replete. Sometimes it will swing out of the nourishing spray before it is full, pulled away by the accumulating weight of water in the other trumpets. At other times, it will spin gently to disgorge into the pool below. Sometimes, two trumpets spill their waters almost simultaneously. At others, one performs its climactic act alone before rising to drink again.

Watching the hypnotic effect of the planetary cycles and epicycles is like witnessing a ballet choreographed according to nature's own laws of complexity. Compared to that classic of chaotic motion, the coupled pendulum, the variables are multiple, combining the unpredictability of the water spout — both in itself and as affected by the breeze — with the interactive ensemble of the rotating components. The two rotational ensembles work according to the same principles, as we can intuitively sense, yet they never perfectly repeat the same phase, either in themselves or in concert.

The design is uncompromisingly abstract. Yet there is an elusive feeling of things remembered from nature — of fish rising to the surface to gulp air, of elegant water birds drinking at a waterfall, of flowers dipping their heads in the wind and rain. Through his artistic

insights and constructional skill, Shingu is involved in a quest to ensnare the patterns underlying the variousness of nature in motion. Many a scientist will share the motivation behind this quest.

Martin Kemp is in the Department of the History of Art, University of Oxford, Oxford OX1 2BE, UK.



Susumu Shingu's *Locus of Rain*, located at Fiat's Lingotto factory in Turin, Italy.

Burnt into memory

Thomas Elbert and Maggie Schauer

As you read this, 35 million humans in various parts of the world are fleeing from war. Their daily lives are severely affected by the psychological consequences of traumatic stress. Today's military actions no longer resemble those of wars long ago, in which one country's army fought the governmental forces of another. In today's wars, more than 80% of casualties are civilians.

A civilian's homeland and life are drastically altered by war. Wartime strategies are ruled by hate and exclusion on both sides, with an attempt to unify one's own group through crimes against humanity. Wars are accompanied by systematic killing and ethnic cleansing — regions are left uninhabitable for local people, landmines remain a constant threat, and cultural heritage and monuments are destroyed. During the Kosovo crisis, one 90-year-old woman brought the crux of modern warfare into focus: "I've lived through the First and Second World Wars but this was worse. This time it was so bad that even the cows ran away. In the night of 24 March, at 3:30 in the morning, as NATO bombs began falling over Yugoslavia, I saw black-masked paramilitia running through Djakovica, shooting, cutting throats and burning houses."

In our attempts to understand the psychophysiological consequences of these atrocities, we have worked with war victims from crisis regions such as the Balkans, the West Nile and Somalia, conducting interviews, observing behaviour and measuring physiological responses to specific stimuli. We are still amazed at the ability of an illiterate survivor, who has been driven out of the bush in southern Sudan and who has had little contact with the outside world, to present us with a classic report of textbook psychiatric symptoms.

Post-traumatic stress disorder (PTSD) is

defined by a specific pattern of core symptoms: re-experiencing intrusions through nightmares and flashbacks — moments of recollection so intense that the victims believe themselves to be back amid the atrocities; an exaggerated startle response and sustained preparedness for an instant alarm response (hyperarousal); difficulty in calming down or falling sleep (describing a readiness for flight or fight, rather than a permanently enhanced autonomic activation); and active avoidance of places where danger was previously experienced, and/or passive avoidance marked by an avoidance of thoughts or feelings related to the traumatizing event. In severe cases, the symptoms may also include dissociation, de-realization, depersonalization or persecutory delusions.

Even though PTSD seems to be present in all corners of the globe, some argue that the disease is a Western concept, and disagree with the empirical testing of this diagnosis for reasons such as a 'medicalization' of politics or neocolonialism. Often, psychosocial-aid organizations believe that scientific investigations intrude on non-Western cultures and should not be carried out there. However, it is clear from scientific observations that PTSD and its symptoms are present across cultures, with the only differences being the culturally specific expression of symptoms and the indigenous ways in which sufferers deal with them. Studies have shown that cross-cultural similarities and consistencies greatly outweigh cultural and ethnic differences. There must be a common underlying basis for these symptoms. We think that psychophysiology, and the structure of memories in particular, provides the answers.

To understand the physiological basis of PTSD and the common thread that underpins it, the concept of memory must be understood in relation to trauma. Autobiographical memory can be divided into 'cold' memories, which include knowledge about periods of our lives and specific events, and 'hot' ones, which comprise emotional and sensory memories. Cold memories (for example, "on 24 March at 3:30 I was living on my farm in Djakovica; we had three cows") are usually connected with hot sensory memories, ("black-masked, dark night, shooting, burning houses") as well as with cognitive ("I can't do anything"), emotional (fear, sadness) and physiological (heart racing, fast breathing, sweating) elements. It is through this string of hot memories that a network of fear is constructed.

Physiological changes that take place in the brains and bodies of survivors of organized violence have been shown to be affected by traumatic stressors and are linked to the

Psychological trauma

A cut into the soul as a result of a horrifying experience can persist as a crippling disease with its core conceptualized as post-traumatic stress disorder.

organization of memory. These changes may include triggering of stress-related systems (such as the hypothalamus–pituitary–adrenal cortex axis), or changes in the functioning and even the structure of the medial temporal lobe and connected limbic networks in the brain.

When an organism is driven down the defence cascade — when flight is impossible, fight futile and only a startling freeze is left in our evolutionary repertoire — the functioning of the medial temporal lobe structures that are the portal for autobiographical memory is altered such that hot and cold memories lose their interconnectivity. We hypothesize that it is this disconnection that causes those who experience flashbacks to become entangled in fear and anxiety. The victim is unable to locate the flashback in time and space because the hot memory is not connected to the cold memory. If this connection could be restored, the horror and 'reality' of the emotions associated with the traumatic memory might be alleviated.

In the long run, the wound that the mind has sustained cannot be entirely healed. But there are approaches that can help. First, reweaving the contents of hot memories back into cold-memory networks can bring relief from the burns of psychological trauma. Second, documenting and acknowledging human-rights violations can dignify the hot traces left in the memory of those who have survived terror and organized violence. ■

Thomas Elbert and Maggie Schauer are in the Department of Psychology, University of Konstanz, 78457 Konstanz, Germany, and at Vivo, Cupramontana, Box 17, I-60032 Ancona, Italy.

FURTHER READING

- Friedman, M. J. *Post Traumatic Stress Disorder* (Compact Clinicals, Kansas City, 2001).
 Lang, P. in *Emotions: Essays on Emotion Theory* (eds van Goozen, S. H. M., van de Poll, N. E. & Sergeant, J. A.) 61–93 (Erlbaum, New York, 1994).
 Kaldor, M. *New and Old Wars: Organized Violence in a Global Era* (Polity, Cambridge, 1999).
 Metcalfe, J. & Jacobs, W. J. *PTSD Res. Quart.* **7**, 1–8 (1996).
 Neuner, F., Schauer, M., Roth, W. T. & Elbert, T. *Behav. Cogn. Psychother.* **30**, 205–209 (2002).
 ◆ www.vivo.org



When it pays to waggle

Fred C. Dyer

The wagging dance of honeybees conveys navigational information about where food is to be found. But it seems that the information is valuable only in certain circumstances.

Almost 60 years ago, Karl von Frisch discovered that worker honeybees, after returning home from rich sources of nectar or pollen, can encode the direction and distance of food sources into body movements that he called dances¹ (Fig. 1). Von Frisch argued that spatial information in the dance, together with odours carried by the dancer, could be used by nest-mates to find the food. The discovery of this so-called dance language provoked intensive research² and no small amount of controversy³. It remains a lively area of research, the latest contribution coming from Sherman and Visscher on page 920 of this issue⁴.

A matter that has never been satisfactorily resolved concerns the precise way in which spatial information in the dance increases the ability of the colony to accumulate food from the environment. Although dancing is correlated with a build-up of forager recruits at a food source, it is not absolutely necessary for recruitment to occur, especially when the food is close and odour cues are strong². Even when spatial information is present, it is sometimes ignored in favour of odour cues by recruits searching for food.

Such observations are hardly evidence that the dance does not communicate spatial information, as some sceptics claim³, but they do raise the question of why that information is sometimes used for recruitment and sometimes not. Sherman and Visscher⁴ address this question, and in doing so provide insight into the adaptive function, or survival value, of the dance language. They have studied colonies of the European bee, *Apis mellifera*, in southern California, to see



Figure 1 A dancing bee, here blurred by rapid wagging and surrounded by forager recruits. (Image courtesy of P. K. Visscher.)

what influence the danced information has on a colony's success. They find that there is an effect, but that it depends on the season.

One standard strategy for assessing the adaptive function of a trait is to prevent or alter its expression, and then to see how the manipulation affects the animal's success in the environment. In their application of this strategy, Sherman and Visscher deprived bee dancers of precise directional information, and observed how this affected the recruitment of foragers and the rate of accumulation of food by colonies.

The directional information encoded in the dance is the angle that the bee has flown relative to the Sun's azimuth to reach the food. In the dance, the bee repeatedly runs in a straight line over the comb while vigorously wagging her body. The flight direction is

encoded in the orientation of these wagging runs relative to stimuli present in the nest. If the dancer can see the Sun or another bright light source, she simply aligns herself to match her view of the light with her view of the Sun during the flight. If a light source is not visible but the bee is dancing on a vertical surface (as is normal in the darkness of the nest), then she can align her body to gravity to match the angle flown relative to the Sun. If one deprives the dancers of both gravity and a directional light cue, by rotating the comb to the horizontal and then keeping it in darkness or diffuse light, the dances are typically disoriented^{1,2}.

Sherman and Visscher carried out two experiments to assess the consequences of depriving dancers of directional information in this way. First, they compared the rate of forager recruitment to an artificial flower when dances were oriented or disoriented, using paired colonies and a reciprocal design to control for temporal and colony-level differences. As in earlier experiments², recruitment rates were far lower when dances lacked directional information. That some recruits found the food when dances were disoriented can be attributed to the use of odours.

This experiment adds to the sizeable body of evidence that spatial information in the dance helps recruits to find food². But does this communication system make any difference to the ability of a colony's foraging force to amass resources? The second experiment addressed this question by comparing the rate of gain of mass (presumably, mostly nectar)

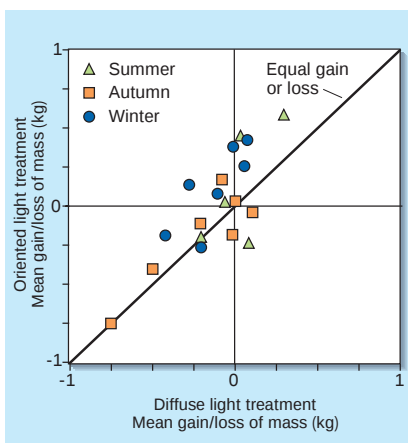


Figure 2 Pointing the way. The figure compares mass gain by honeybee colonies with and without the orientation cues needed for forager recruitment dances to encode the direction of food. The comb on which the dances took place was horizontal, so that gravity could not provide a reference for orientation. Colonies were provided with either diffuse light, in which case dances were disoriented, or a directional light source, in which case dances were oriented. Each symbol corresponds to an observation period during which two colonies were subjected to each treatment. Sherman and Visscher⁴ found that during the winter, but not summer or autumn, the mean mass gain was higher overall for the colonies with the oriented dances than for those with disoriented dances. (Data from ref. 4.)

by colonies during periods of oriented and disoriented dancing. Somewhat surprisingly, the rate of gain was not always higher during periods of oriented dancing. Although there was a strong overall effect of oriented dances, the advantage was pronounced only in the winter (bees can fly throughout the year in southern California). There was no advantage associated with oriented dances during the summer or autumn (Fig. 2).

Sherman and Visscher attribute these patterns to seasonal variation in the distribution of local resources. Oriented dances made no difference during the autumn dearth of nectar, when all colonies lost mass overall, or during the summer period of nectar abundance, when colonies easily gained mass. Oriented dances did make a difference, however, during winter, when rich floral resources were available, but were more dispersed in space and time than during the summer. This result leads to an interpretation of the adaptive function of the dance language that is more precise than is usually supposed. Rather than providing a general advantage in the social exploitation of floral resources⁵, it appears that the spatial information in the dance offers a specific advantage when resources are patchy and ephemeral. Under such conditions, when the probability of survival might otherwise be marginal, the dance language may tip the balance in favour of colony growth.

This study⁴ paves the way for a more detailed examination of the ecological circumstances in which spatial communication increases the success of colonies in

gathering food. In particular, the authors' speculations about the natural selection pressures that might have shaped the dance during its evolution in the tropical forests of Asia could be explored through field studies that compare the reliance on spatial communication by different honeybee species. All *Apis* species other than the European bee studied by Sherman and Visscher live in Asia, where honeybees presumably first evolved⁶. More generally, it would be interesting to examine the adaptive function of spatial communication in a broader context: other social insects, especially ants⁷ and stingless bees⁸, also have ways of communicating the precise spatial location of food, but these primarily involve odour trails. Sherman and Visscher's general approach provides a model for investigating the specific circumstances under which spatial communication enhances the biological competitiveness of these organisms. ■

Fred C. Dyer is in the Department of Zoology, Michigan State University, East Lansing, Michigan 48824, USA.

e-mail: fcdyer@msu.edu

1. von Frisch, K. *The Dance Language and Orientation of Honeybees* (Harvard Univ. Press, Cambridge, Massachusetts, 1967).
2. Dyer, F. C. *Annu. Rev. Entomol.* **47**, 917–949 (2002).
3. Wenner, A. M. & Johnson, D. L. *Anatomy of a Controversy* (Columbia Univ. Press, New York, 1990).
4. Sherman, G. & Visscher, P. K. *Nature* **419**, 920–922 (2002).
5. Seeley, T. D. *The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies* (Harvard Univ. Press, Cambridge, Massachusetts, 1995).
6. Ruttner, F. *Biogeography and Taxonomy of Honeybees* (Springer, Berlin, 1988).
7. Hölldobler, B. & Wilson, E. O. *The Ants* (Harvard Univ. Press, Cambridge, Massachusetts, 1990).
8. Nieh, J. C. *Am. Sci.* **87**, 428–435 (1999).

with metallicities less than 1/100 that of the Sun, but none with convincingly less than 1/10,000 of the solar metallicity^{3,5}.

The apparent absence of any stars in the Galaxy today with such low metallicities has been interpreted as evidence that such stars never formed³. The process of star formation from gas with low metal content is thought to be different from star formation in today's Universe. With no metals to radiate away the heat generated by the contraction of the gas cloud, star formation is less efficient. Moreover, theory indicates that in metal-poor gas the formation of high-mass, not low-mass, stars is favoured³. Evidence for this is seen in excesses of the so-called 'alpha' elements — oxygen, magnesium and silicon — in metal-poor stars, elements produced primarily during the evolution of massive stars. These early, massive stars would have evolved quickly, leaving the metals they produced and dispersed to the interstellar medium through supernovae explosions as the only evidence that they ever existed. No longer-lived, lower-mass siblings would remain to the present day.

But the barrier at 1/10,000 of the solar metal abundance now seems to have been breached. A new spectroscopic survey of the southern sky, reaching fainter stars than earlier surveys and covering ten times as large a volume of sky, identified the star HE0107–5240 as possibly being extremely metal-poor. Spectroscopic follow-up observations by Christlieb *et al.*⁴ at the European Southern Observatory's Very Large Telescope have confirmed the low metallicity, now measured as 1/200,000 of the solar metal abundance — a fraction 20 times less than in the previously known, most metal-poor stars. Even the strongest spectral lines of metals are barely visible in its spectrum.

The discovery of HE0107–5240, at a distance of half the diameter of the Galaxy from the Sun and with a mass four-fifths that of the Sun, shows that low-mass, metal-poor stars could and did form in the early Galaxy, despite theoretical indications to the contrary. The chemical composition of HE0107–5240 may provide a clue to the mystery of its formation. Although extremely deficient in metals, the star displays a striking excess of the light elements carbon and nitrogen. There is 10,000 times more carbon and 200 times more nitrogen relative to iron. These huge excesses of carbon and nitrogen may have been deposited on the star by a more massive companion star that has long since evolved and disappeared.

In fact, a large proportion of stars with metal deficiencies 1/10,000 that of the Sun show excesses of carbon and/or nitrogen, and some are known to be binaries⁶. Perhaps most, if not all, of the extremely metal-poor stars in the Galaxy today originally formed as part of a binary pair, a by-product of the formation of a massive star. Alternatively,

Astronomy

Relic of the dawn of time

Catherine A. Pilachowski

Elements heavier than helium are synthesized in stars. But could there be stars, created soon after the Big Bang, that contain almost no heavy elements? The discovery of such a star gives new clues to this early time.

For half a century, astronomers have known that the elements of the periodic table are produced in stars¹. The chemical enrichment of the Galaxy has proceeded from an early state with very low amounts of 'metals' (elements heavier than helium) to the present epoch, in which metals make up some 2–3% of the visible mass of the Milky Way². This realization led to the quest for 'population III' — stars containing no metals, or, at least, such a small quantity of heavy elements as to be insignificant. These stars³ would be relics from very early in Galactic history, before star formation and evolution polluted the primordial gas with metals. After decades of searching, such a star may have been found, as Christlieb *et al.*⁴ report on page 904 of this issue.

Stars that contain very small amounts of heavy elements are rare in the Galaxy. Among the brightest 10,000 stars in the sky, only a handful contain really low amounts of metals. These metal-poor stars belong to the Galactic halo, that faint sphere of old stars and globular clusters in which the Galactic disk is embedded (Fig. 1). The orbits of halo stars are eccentric, taking the stars through the Galactic disk at relatively high velocity. Searches for metal-poor stars are carried out in two ways: by identifying high-velocity stars that are confirmed as halo objects by the details of their motion; and by finding stars whose radiation bears the spectral signatures of very low metal abundances. Extensive searches for stars of low metal abundance over the past two decades have found many halo stars

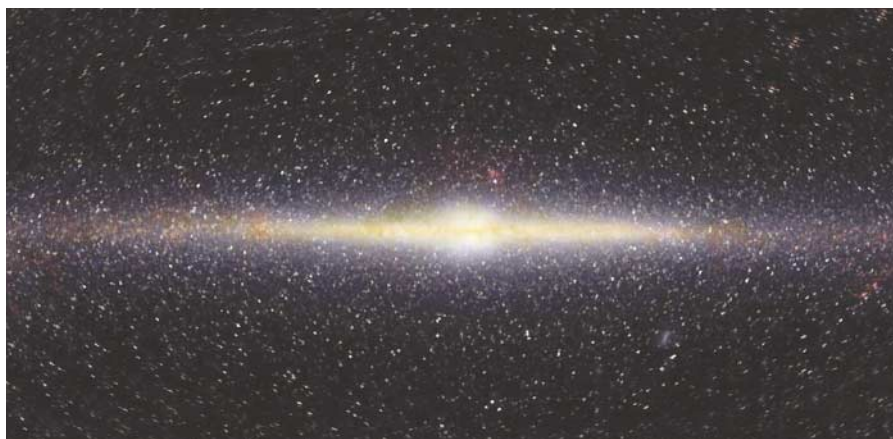


Figure 1 The Galactic disk. This infrared image of the Milky Way was taken by the COBE satellite. Christlieb *et al.*⁴ have discovered a star in the Milky Way whose heavy-element content is extremely low. This star could belong to 'population III' — a family of stars that formed much earlier in the history of the Universe than other existing stars, as little as 1 billion years after the Big Bang.

the high carbon and nitrogen abundances could be a result of extensive internal mixing, which may occur in stars of such low metallicity.

Is HE0107–5240 just the first of many metal-poor stars to be found as the reach of surveys extends farther out from the Galaxy's inner halo? Will stars be found that fill the metallicity gap between 1/10,000 and 1/100,000 of the solar metal abundance, tracing the rise in Galactic metal abundance smoothly from earliest times to the present day? The previous 1/10,000 limit could be an artefact of the relatively small volume of Galactic space explored by earlier surveys. Or does HE0107–5240 represent a different epoch of star formation entirely — perhaps, as suggested by Christlieb *et al.*⁴, an era before the 'reionization' period of

the Universe's history, just 1 billion years after the Big Bang? Only time, and the extension of new surveys to ever fainter stars, will tell.

Catherine A. Pilachowski is in the Department of Astronomy, Indiana University, Bloomington, Indiana 47405, USA.

e-mail: catyp@astro.indiana.edu

1. Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. *Rev. Mod. Phys.* **29**, 547–650 (1957).
2. Grevesse, N. & Anders, E. in *Solar Interior and Atmosphere* (eds Cox, A. N., Livingston, W. C. & Shapley, M. S.) 1227–1234 (Univ. Arizona Press, Tucson, 1991).
3. Bond, H. E. *Astrophys. J.* **248**, 606–611 (1981).
4. Christlieb, N. *et al. Nature* **419**, 904–906 (2002).
5. Beers, T. C. in *The Third Stromlo Symposium: The Galactic Halo* (eds Gibson, B. K., Axelrod, T. S. & Putman, M. E.) *Astron. Soc. Pacif. Conf. Ser.* **165**, 202–212 (1999).
6. Spite, M. *et al. Astron. Astrophys.* **360**, 1077–1085 (2000).

Materials science

Nanomaterial advantage

Ruslan Valiev

Materials may be strong or ductile, but rarely both at once. The processing of copper into a nanostructure possessing different-sized grains produces a material that retains its high strength and ductility under deformation.

Nanocrystalline materials¹ are three-dimensional solids composed of nanometre-sized grains, or crystallites. Because of their unique structure, which is characterized by ultrafine grains and a rather high density of crystal-lattice defects, these materials have extraordinary fundamental properties that could be exploited to make 'next-generation' superstrong metals, ductile ceramics and wear-free materials. On page 912 of this issue, Wang *et al.*² report their success in engineering an unusual nanocrystalline material that combines two useful properties that are often mutually exclusive — strength and ductility.

Strength and ductility are the central mechanical properties of any material. They are determined by the physical nature of plastic deformation, which in conventional, coarse-grained metals is mainly carried by dislocations — line defects of the regular crystal lattice — within individual grains. Models suggest^{3–5}, however, that the mechanism of plastic deformation in nanocrystalline materials may be different and so may lead to novel mechanical properties.

High strength, or hardness, has already been observed in many nanomaterials. But in most cases such nanomaterials have very low ductility — they fail when their shape

is changed. Some even become brittle when a force or deforming load is applied^{6,7}. Strength and ductility are usually opposing characteristics: the higher the strength, the lower the ductility, and vice versa. This correlation is associated with the nature of plasticity: the more difficult it is for dislocations to appear and to move, the stronger but more brittle and less ductile is any crystalline material.

Wang *et al.*² have taken an original approach to this problem. They created a copper nanostructure by rolling the metal at low temperature — below 77 K, the temperature of liquid nitrogen — and then heating it up to around 450 K. The result was a 'bimodal' structure of micrometre-sized grains (at a volume fraction of around 25%) embedded in a matrix of nanocrystalline grains. The material showed extraordinarily high ductility, but also retained its high strength. The reason for this behaviour, the authors believe, is that, while the nanocrystalline grains provide strength, the embedded larger grains stabilize the tensile deformation of the material. New mechanisms are involved here, including an unusual twinning mode in which crystal domains move in a concerted way.

An advantage of Wang and colleagues' nanocrystalline material over many other nanomaterials is the way in which it was produced — through heavy plastic deformation by rolling. Unlike nanomaterials produced by the compacting of powders, nano-structured metals made by severe plastic deformation do not have any residual porosity or contaminations⁸. Their mechanical behaviour is an integral property of the material itself, and is not determined by the processing technique used.

Wang *et al.*² studied the properties of copper, but can their approach be applied to other metallic nanomaterials? The authors argue that their technique is universal, and this is supported by the results of computer simulations in which they modelled the mechanical behaviour of other metals possessing a bimodal structure. It seems reasonable to expect that severe plastic deformation and thermal treatment could certainly be applied to various other metals and alloys with the aim of producing a bimodal structure, and the experimental demonstration of other nanomaterials with high strength and ductility would be an interesting task.

In fact, another approach has been put forward by my colleagues and me which also leads to high strength and ductility in nanostructured materials⁹. Here, the idea is to change the deformation mechanism and induce so-called rotational plasticity, which is associated with the rotation of grains in the material. This kind of deformation becomes possible through the formation of a fine-grained structure on the scale of hundreds of nanometres, with predominantly high-

angle grain boundaries at which sliding can occur.

Investigating the extraordinary strength and ductility of nanostructured materials is of fundamental as well as practical importance. At the fundamental level, the insight into deformation mechanisms improves our understanding of mechanical behaviour. On the practical side, there is the prospect of using nanostructured metals and alloys for many applications, among them micromechanical systems and biomedical implants. In fact, processing nanomaterials to achieve both strength and ductility can enhance the material's resistance to fatigue¹⁰. More than anything else, it is fatigue that at present limits the lifetime, and hence the range of applications, of many advanced materials. ■

Ruslan Valiev is at the Institute of Physics of Advanced Materials, Ufa State Aviation Technical University, Ufa 450000, Russia.
e-mail: RZValiev@mail.rb.ru

1. Gleiter H. *Acta Mater.* **48**, 1–29 (2000).
2. Wang, Y. M., Chen, M. W., Zhou, F. H. & Ma, E. *Nature* **419**, 912–915 (2002).
3. Ovid'ko, I. A. *Science* **295**, 2386 (2002).
4. Van Swygenhoven, H. *Science* **296**, 66–67 (2002).
5. Yamakov, V., Wolf, D., Phillpot, S. R., Mukherjee, A. K. & Gleiter, H. *Nature Mater.* **1**, 45–49 (2002).
6. Koch, C. C., Morris, D. G., Lu, K. & Inoue, A. *Mater. Res. Soc. Bull.* **24**, 54–58 (1999).
7. Morris, D. G. *Mechanical Behaviour of Nanostructured Materials* (Trans. Tech. Publ., Uetikon/Zurich, 1998).
8. Valiev, R. Z., Islamgaliev, R. K. & Alexandrov, I. V. *Prog. Mater. Sci.* **45**, 103–189 (2000).
9. Valiev, R. Z., Alexandrov, I. V., Zhu, Y. T. & Lowe, T. C. *J. Mater. Res.* **17**, 5–8 (2002).
10. Hoepfel, H. W. & Valiev, R. Z. *Z. Metallkunde* **93**, 641–648 (2002).

Chemistry

Only skin-deep

Peter J. Rossky

Separating the details of surface structure from those of the bulk material is difficult. An X-ray-based technique reveals clearer, and contrasting, pictures of the atomic surface structure of water and methanol.

A target in modern chemical research is the development of a truly atomistic description of complex chemical processes. An enormous number of processes occur in bulk solutions, but many occur at liquid interfaces, such as those at biological membranes, in emulsions and at the surface of droplets. A surface need not have the same structure as the bulk material that lies underneath. Atoms at surfaces are subject to unbalanced forces and so can often rearrange to a new more stable structure. Such surface reconstruction is well known, for example, at the surface of solid silicon, and occurs at macroscopic liquid surfaces as well. But it is difficult to see the atoms at the surface of a liquid without seeing the atoms of the 'ordinary' liquid below. So the direct observation of the atomic-level structure of the surface of a liquid is a challenge — one that has now been met by Wilson *et al.*¹, writing in the *Journal of Chemical Physics*. They describe the first direct measurement of the spacing between molecules at the surface of liquid water and of liquid methanol.

The chemical behaviour of liquids at their surface is of special importance for the way in which gases and liquids interact, and considerable research into these processes has been conducted². Examples include the adsorption of oxygen from the air in our lungs and the industrial process of 'scrubbing' waste gas. A further, environmentally critical, case is the chemistry of water droplets in the Earth's atmosphere. It is also worth noting that a vacuum is the ultimate non-polar, and therefore hydrophobic, sur-

face, so that understanding the free surface of an aqueous solution is instructive in this respect as well.

Ways of determining the spacing between molecules in bulk materials are relatively well established. The familiar crystal structures of solids are typically inferred from the pattern of diffracted X-rays scattered from the atoms. Atomic distributions in bulk liquids can be obtained in a similar way. But for interfaces, such as a liquid surface, 'surface selectivity' is required. At solid surfaces, this is straightforward: the system is rigid and the space above it can be evacuated (so that no molecules lie between the observer and the surface); then probing by scattering of light or particle beams from the surface can provide direct structural information.

For liquids, the situation is more complicated, but there are several avenues of investigation. Molecular simulation has historically provided the most detailed atomic-level views of liquid interfaces³ (Fig. 1), but direct experimental probes are complementary and are increasingly providing structural views and critical tests of theory. For macroscopic surfaces of liquids, and for interfaces between liquids or between liquids and solids, scattering methods can provide considerable information about interfacial densities^{2,4,5}. Details of molecular orientation and vibrational motion come from methods that detect light that interacts only with molecules present in the asymmetrical environment of an interface⁶. But for relatively volatile liquids, such as water and alcohols, a direct view of the atomic positions on

the free liquid surface is complicated, in part by the presence of the vapour.

This is a problem that Wilson *et al.*¹ address to directly observe local atomic structure. To combat volatility, they used a jet of liquid, just micrometres in diameter, passing through a small chamber so that vapour could be pumped away. Then, to explore the surface structure of the liquid in the jet, they used a technique known as EXAFS — extended X-ray absorption fine structure. EXAFS exploits variations in the X-ray-absorbing power of the liquid with X-ray wavelength, which depend on the type, number and position of atoms surrounding the atom that is absorbing the X-rays. To obtain a view of surface molecules only, the authors identified the ions emerging from the liquid surface, generated by X-ray absorption in that region.

Water has the unusual ability to form a three-dimensional network of hydrogen bonds in the bulk material. In contrast, methanol, with only one polar hydrogen atom to donate in a hydrogen bond, tends to form linear hydrogen-bonded chains in its liquid form. Minimizing the number of unsatisfied 'dangling' hydrogen bonds at a surface is the force that drives the reconstruction of the liquid surface in both cases. But the presence of a vapour interface necessarily interrupts some of the water hydrogen bonding. For methanol, as Wilson *et al.*¹ discuss, measurements of vibrational frequencies had suggested that methanol surface reconstruction involves the formation of a molecular bilayer on the surface. The molecules tend to be oriented so that the upper and lower bilayer surfaces are methyl

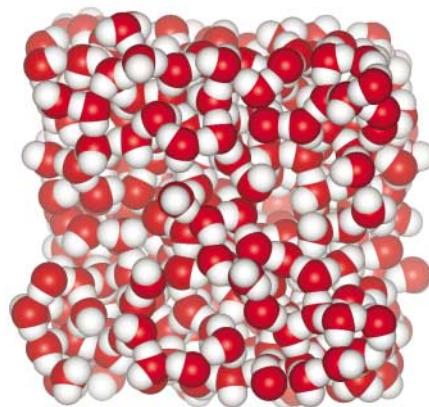


Figure 1 Skimming the surface. This image of an instantaneous configuration of the surface of liquid water at a vapour interface is the result of molecular simulation (oxygen atoms are red, hydrogen white). Look carefully and the presence of unsatisfied hydrogen bonds for molecules at the surface is clear. X-ray measurements by Wilson *et al.*¹ now confirm the picture, showing that the decrease in organized hydrogen bonding at the water surface results in increased intermolecular spacing — but that the reverse is true for methanol. (Image courtesy of Sandro R. P. da Rocha, Wayne State University.)

groups and the hydrogen-bonding OH groups are sandwiched in between. So in this case, the surface reconstruction need not interrupt the linear chains of hydrogen bonds formed within the 'sandwich'.

The observations made by Wilson *et al.* show directly that the intermolecular separation in water at the surface is about 5% greater than in the bulk liquid, and so is similar to that of a gas-phase dimer. At the surface of methanol, however, the separation is about 5% lower, reaching a value close to that for the crystalline solid. In water, the increasingly organized hydrogen-bonded structure, from dimer to liquid to ice, leads to increasingly stronger attractive forces and increasingly shorter intermolecular distances. The results of the EXAFS study reflect the decrease in organized hydrogen bonding at the water surface and the increase in order at the surface of liquid methanol.

The broader impact of this new EXAFS experiment will depend on how well it can be generalized to describe solutions. Among the implications for interfacial chemistry, the direct observation of the solvation of surface-adsorbed species is a potential outcome. The method measures the average surroundings of all atoms of the same elemental type, so a new focus must be on systems in which the solute has a distinguished

atom, of different elemental type. Ionic solutions are a practical example in atmospheric chemistry, as are solutes with distinguished atoms such as chlorine.

Although interatomic distances have a distinct signature in EXAFS, the composition of the local environment cannot at present be well determined in this surface case. One need, and challenge, is a new theoretical description of EXAFS signals in regions of inhomogeneous density — density that at a liquid surface varies very rapidly indeed, from liquid to vacuum in less than a nanometre. It is clear that the various approaches to interfacial characterization will continue to be complementary. A synergistic element that should not be underestimated is that the present work and further applications will provide a benchmark for direct tests of the molecular models that underpin atomistic simulation.

Peter J. Rossky is in the Department of Chemistry and Biochemistry, University of Texas, Austin, Texas 78712-1167, USA.

e-mail: rossky@mail.utexas.edu

1. Wilson, K. R. *et al.* *J. Chem. Phys.* **117**, 7738–7744 (2002).
2. Nathanson, G. M., Davidovits, P., Worsnop, D. & Kolb, C. E. *J. Phys. Chem.* **100**, 13007–13020 (1996).
3. Benjamin, I. *Acc. Chem. Res.* **28**, 233–239 (1995).
4. Schlossman, M. *Curr. Opin. Coll. Int. Sci.* **7**, 235–243 (2002).
5. Penfold, J. *Rep. Prog. Phys.* **64**, 777–814 (2001).
6. Richmond, G. *Annu. Rev. Phys. Chem.* **52**, 357–389 (2001).

Molecular biology

RNA gets a grip on translation

Jack W. Szostak

These are interesting times for those who study RNA. The latest curious discovery is that some messenger RNAs have sequences that sense small molecules directly, so controlling translation of the RNA into protein.

Over the past decade, researchers have applied the principles of darwinian evolution in the laboratory to generate hundreds of RNA sequences that fold into unique shapes and thereby bind to specific target molecules. The experiments revealed the capacity of these folded RNAs — commonly known as aptamers — to recognize virtually any molecular target, but led to a growing mystery: why is this ability not used in nature? Two papers from a group led by Ron Breaker, published on page 952 of this issue¹ and in *Chemistry and Biology*², now dissolve this conundrum, by showing that regulatory sequences in certain bacterial messenger RNAs can in fact sense small molecules directly.

Like most biological processes, the biosynthesis and import of vitamins B₁ (thiamine and its active form, thiamine pyrophosphate), B₁₂ (adenosyl-cobalamin) and B₂ (riboflavin) by bacteria requires specific enzymatic and transport proteins, which are produced from genes via messenger RNA

intermediates. These mRNAs contain protein-coding stretches of sequence, as well as non-protein-coding stretches^{3–5}. The non-coding sequences are needed to control translation of the mRNAs into proteins: they are part of signal-transduction pathways that sense the level of a given metabolite (the relevant vitamin), and use that information to control the efficiency of enzyme production.

Normally (that is, in all examples known until now), control regions such as these are bound by metabolite-sensing proteins, which regulate the conformation of the mRNA and thus the access of the ribosome — the protein-making machine — to translation-initiation signatures within the mRNA. Such regulatory proteins had been sought for bacterial vitamin biosynthesis, but had not been found. So it was proposed that in these cases the RNA might interact directly with its cognate metabolite^{3–5} (reviewed in ref. 6). Support for this hypothesis was particularly strong in the case of vita-

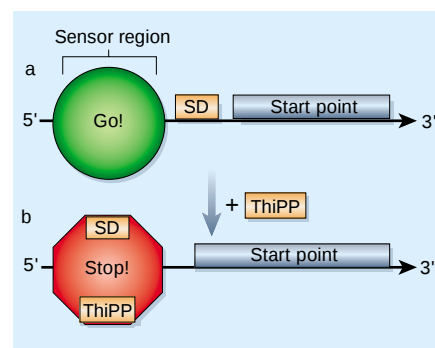


Figure 1 RNA as metabolite sensor and translation regulator. The figure represents part of a messenger RNA that encodes a bacterial protein involved in the biosynthesis of vitamin B₁ (thiamine pyrophosphate, ThiPP). This part contains a region involved in sensing ThiPP; the so-called Shine–Dalgarno (SD) sequence, which is recognized by the ribosome; and a sequence that marks where the enzyme-encoding portion of the mRNA begins. a, Winkler *et al.*¹ have found that in the absence of ThiPP, the sensor region adopts a conformation that exposes the Shine–Dalgarno sequence. This would allow the ribosome to bind and begin translation. b, Binding to ThiPP causes the sensor region to change shape, obscuring the Shine–Dalgarno sequence. Thus ThiPP controls the production of one of its biosynthetic enzymes directly via a sequence within the enzyme-encoding mRNA.

min B₁₂, which could inhibit ribosome binding to a relevant mRNA in a purified system containing just ribosomes and mRNA.

To obtain more direct evidence for an RNA–metabolite interaction, Winkler *et al.*¹ and Nahvi *et al.*² synthesized the previously defined regulatory regions of two mRNAs coding for enzymes involved in B₁ biosynthesis, and one mRNA that codes for a B₁₂ transporter, respectively. The authors proved that a direct interaction occurred by observing changes in the pattern of spontaneous, hydroxide-ion-mediated cleavage of the mRNAs in the presence of the target vitamin. Independent confirmation of RNA–metabolite binding was provided by equilibrium dialysis.

The patterns of spontaneous cleavage events also helped the authors to define the secondary structures — the patterns of base-pairing between segments of the RNA — of the active, folded RNA sequences. The vitamin-dependent changes in these patterns pointed to a conformational alteration in the folded state of the RNAs upon vitamin binding. These data showed that a translation-initiation signature known as the Shine–Dalgarno ribosome-binding site was accessible in the unbound state, but became occluded in the bound state (Fig. 1). Winkler *et al.* and Nahvi *et al.* then generated a series of mutant RNAs that were expected to favour one of these two states of the Shine–Dalgarno

sequence. The results strongly supported the hypothesized vitamin-controlled structural transition, explaining the observed control of mRNA translation. Similar evidence has been claimed (but not yet published) for an mRNA control region required in vitamin B₂ biosynthesis.

One puzzle is that the vitamin-sensing mRNA sequences that control translation^{1,2} are considerably larger, at up to 200 nucleotides, than the aptamers that have been generated *in vitro*, which typically consist of less than 40 nucleotides. Why is this? The most likely explanation is that the mRNA control regions do much more than just bind a target molecule. They function by virtue of their ability to fold into two distinct conformations, which are finely balanced energetically so that a change in vitamin concentration drives the transition from one conformation, which allows translation to begin, to the other, which does not. However, several groups, including Breaker's, have 'evolved' RNA elements *in vitro* that can switch between conformations in a ligand-dependent way, yet are much smaller than the *in vivo* examples. Perhaps the large size of the *in vivo* regulatory sequences is an 'in vivo artefact' of the constraints on the evolution of functional RNA structures from a limited set of initial sequences. In contrast, the *in vitro* situation generally

involves sampling from a large set of random sequences.

Why have metabolite-sensing RNA sequences been found (so far) only in mRNAs involved in vitamin biosynthesis and import? One intriguing possibility, suggested by Winkler *et al.*¹, is that these RNA control elements are ancient, dating from the 'RNA world' — a hypothesized early stage in the evolution of life on Earth, when proteins did not exist. All three vitamins in question are biochemically more or less RNA-like (B₁₂ contains an adenosine group), and have themselves been proposed to date from the RNA world⁷. But many other metabolites, and their biosynthetic pathways, must be as old or older. With so much of biochemistry now subject to the domination of proteins, it is unclear why the biosynthesis of these particular metabolites should have remained so staunchly within the realm of RNA. Perhaps these issues will be clarified if and when additional examples of RNA-mediated metabolic control are discovered.

Whatever the answer, it is clear that researchers investigating translational control must bear in mind that non-coding mRNA sequences might regulate this process directly. A different example of this phenomenon was provided recently by Johansson *et al.*⁸, who described an mRNA regulatory sequence that acts as a direct tem-

perature sensor in a pathogenic bacterium, *Listeria*, controlling translation without needing any regulatory proteins. This RNA sensor detects the increase in temperature that occurs when the bacterium moves into a mammalian host, and regulates the expression of genes associated with bacterial virulence. Along with the discovery of the numerous, widespread micro-RNA molecules, which seem to regulate the translation of mRNAs in a variety of ways, these findings provide striking examples of new and unexpected roles for RNA in controlling gene expression. Given the current pace of discovery, it seems likely that yet more surprises may be just around the corner. ■

Jack W. Szostak is at the Howard Hughes Medical Institute, Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.

e-mail: szostak@molbio.mgh.harvard.edu

1. Winkler, W., Nahvi, A. & Breaker, R. R. *Nature* **419**, 952–956 (2002); advance online publication, 16 October 2002 (doi:10.1038/nature01145).
2. Nahvi, A. *et al. Chem. Biol.* **9**, 1043–1049 (2002).
3. Nou, X. & Kadner, R. J. *Proc. Natl Acad. Sci. USA* **97**, 9170–9175 (2000).
4. Miranda-Rios, J., Navarro, M. & Soberón, M. *Proc. Natl Acad. Sci. USA* **98**, 9736–9741 (2001).
5. Gelfand, M. S. *et al. Trends Genet.* **15**, 439–442 (1999).
6. Stormo, G. D. & Yongmei, J. *Proc. Natl Acad. Sci. USA* **98**, 9465–9467 (2001).
7. White, H. B. III *J. Mol. Evol.* **7**, 101–104 (1976).
8. Johansson, J. *et al. Cell* **110**, 551–561 (2002).

Environment

Trash trends

Recycling campaigners and environmental groups often quote alarming statistics on how much waste one person produces each year. But a comprehensive report on rubbish collected in New York City throughout the twentieth century claims that the figure has dropped dramatically, from its peak in the 1940s.

Writing in *Environmental Science and Technology* (doi: 10.1021/es011074t), Daniel Walsh describes the rise and fall of garbage components in the Big Apple. Using the most complete set of municipal records for a US city's residential refuse, Walsh records a maximum output of 940 kg of waste per person in 1940, and a low of 320 kg per person in both 1961 and 1963.

Surprisingly, since the 1980s a person's annual throwaways have stabilized at a relatively low 430 kg. Also, the most significant trash trends were mostly declines in the percentages of different categories of waste. The relative amounts of fuel ash, food waste, metal and



glass dropped, but the percentage of plastics rose and that of paper remained the same. The photograph on the left shows ash collection earlier in the twentieth century; that on the right a present-day scene.

Between 1920 and 1990 there was a 50% decrease in refuse density. Walsh attributes this to a decrease in coal and other fuel ash,

and to technologies that have enabled product packaging to switch from glass and metal to paper and plastic. Over the same period, organic-matter waste rose fourfold, increasing the greenhouse-gas potential per unit of dumped or incinerated garbage. Walsh estimates that the totality of New York City's refuse for the past



century represents a carbon pool of 80 million tons.

As coal was at the start of the twentieth century, paper is now the most abundant category of refuse, accounting for roughly 35% of all residential discards. Will the figures for the twenty-first century reflect the long-forecast advent of the paperless office? **Kendall Powell**

Neurobiology

Social eating for stress

Marla B. Sokolowski

One type of the nematode worm *Caenorhabditis elegans* feeds alone, another in aggregates. The neurobiological underpinnings of these behaviours are now being revealed at the molecular level.

Behavioural ecologists have shown that many animals form social groups in response to stressful environmental conditions. Neurobiological evidence for this behaviour has now been discovered in the nematode worm, *Caenorhabditis elegans*. On pages 899 and 925 of this issue, de Bono *et al.*¹ and Coates and de Bono² present striking results on the genetic, molecular and neural mechanisms underlying nematode social feeding. These discoveries provide tantalizing insights into the effects of stress in social groupings.

Caenorhabditis elegans has a fixed quota of 302 neurons, and with the availability of modern genetic and molecular tools it has become a superb model for studying behavioural neurobiology. Many of its neurons are being traced to specific functions that link genes and behaviour through neural circuitry³. As far as the *C. elegans* response to food is concerned, there are two naturally occurring foraging behaviours⁴: individuals of one type feed alone (solitary) whereas the other type feeds in aggregates ('social'; the term is used loosely because it is not yet clear that this is true social behaviour). The two variants are known to arise from a single amino-acid difference in a putative neuropeptide receptor protein (NPR-1), which is related to the mammalian neuropeptide Y receptor family. Solitary feeders show reduced locomotion and high dispersal activity when encountering food (a 'lawn' of bacteria); social feeders, by contrast, continue to move rapidly towards food and eventually aggregate on the borders of the lawn.

De Bono and colleagues screened mutant strains of *C. elegans* for mutations affecting social feeding, and found several that interacted with the *npr-1* gene. Using green fluorescent protein as a marker (Fig. 1), they then looked at where the genes concerned are expressed, and identified the relevant neurons. They also targeted the expression of each gene to subsets of the neurons, to determine where each gene is required. To confirm that these neurons are indeed involved in the animals' social behaviour, they ablated each neuronal cell using a laser or blocked its electrical activity. This combination of approaches allowed the authors to identify two sets of neurons. One set is found in the animal's anterior, and is likely to be responsible for sensing the external environment¹. The other occurs in tissues bordering the structure containing the body fluids.

Presumably, cross-talk between the two systems controls the nematodes' solitary or social response to food.

In the first paper, de Bono *et al.*¹ show that social feeding is induced in a crowded feeding environment, after perception of this condition by the anterior neurons known as ASH and ADL. ASH neurons sense both mechanical and chemical stimuli, whereas ADL neurons are involved in odour avoidance. But more notably, these neurons are both nociceptive — they detect pain or aversive conditions such as crowding or food shortage. Ablating either ASH or ADL neurons prevents aggregation in response to crowding. On this evidence, then, social feeding in *C. elegans* may be a response to stressful conditions.

Solitary feeders aggregate when the *npr-1* gene is deleted, suggesting that *npr-1* represses social feeding. *npr-1* is primarily expressed in the nervous system and is not required during development for repression of social feeding². Mutations in other genes, *osm-9* and *ocr-2*, restore solitary feeding behaviour in *npr-1* mutant animals¹. These genes are predicted to encode components — subunits of a membrane cation channel called TRPV — that are thought to form a sensory transduction channel in several *C. elegans*

sensory neurons. The subunits of TRPV are required in either ASH or ADL neurons, both of which are involved in detecting aversive stimuli to trigger social feeding. Interestingly, *C. elegans osm-9* and *ocr-2* are respectively related to genes encoding the mammalian vanilloid (capsaicin) receptor TRPV1 (VR1), which is implicated in sensing thermal pain, and the TRPV4 channel, which responds to mechanical and osmotic stimuli.

In the second paper², Coates and de Bono implicate a different ion channel, in a different place, in regulating social feeding. This channel is controlled by the signalling molecule cyclic GMP, and is composed of the subunits TAX-2 and TAX-4. The data² again support the idea that NPR-1 inhibits social feeding, and that here TAX-2 and TAX-4 regulate responses to signals that promote aggregation by sending antagonistic signals via the body fluid. From an evolutionary perspective, this pathway might have arisen to regulate the relative numbers of solitary and social animals, as each strategy would have associated costs and benefits depending on the environment. Given that neuropeptide Y is involved in regulating food intake in mammals, an intriguing question here is whether *npr-1* differentially affects the food intake of solitary and social-feeding *C. elegans*, and if so what the consequences are on the animals' fitness.

Each paper presents a model to explain how social feeding is regulated, with one system sensing the external environment and the other involving neurons inside the body cavity. But in neither paper is the attempt made to connect the two. It remains unclear how these molecular and neural



Figure 1 Glow worms — social *C. elegans* displaying expression of green fluorescent protein. (Image courtesy of S. Reichelt and M. de Bono, LMB, Cambridge.)

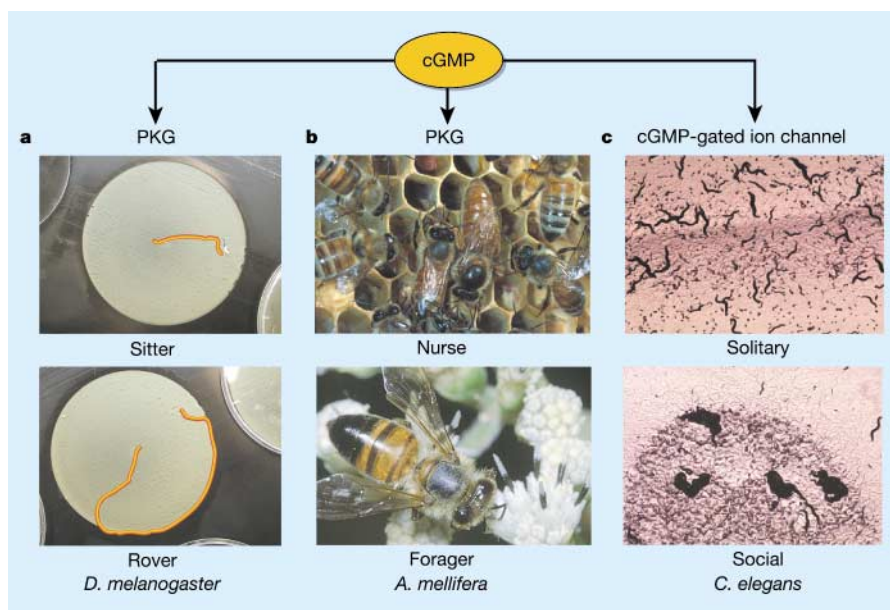


Figure 2 The cyclic-GMP signalling pathway and variation in food-related behaviours. **a**, In the fruitfly *Drosophila melanogaster*, two different food-search behaviours — the self-explanatory ‘sitter’ and ‘rover’ — depend on variants of the *foraging (for)* gene, which encodes a cGMP-dependent protein kinase (PKG)⁵. Shown here are 5-minute foraging trails of fly larvae. **b**, In the honeybee *Apis mellifera*, *for*, and so PKG, is upregulated when bees switch from nursing in the hive to foraging for nectar and pollen⁶. **c**, Aggregation in *Caenorhabditis elegans* involves a cGMP-gated ion channel, as shown by Coates and de Bono². Adverse conditions such as crowding or food shortage affect food-related behaviours in all of these organisms.

mechanisms act in the animal as a whole to regulate social behaviour.

How about the situation in other organisms? Signalling through cGMP is also associated with normal variation in food-related behaviour in the fruitfly *Drosophila melanogaster* and the honeybee *Apis mellifera* (Fig. 2). In two cases, variation in behaviour is determined as natural variation in a single gene — *npr-1* in the nematode, and *foraging (for)* in the fruitfly. The *for* gene encodes a mediator of cGMP action, a protein kinase enzyme known as PKG. As their names suggest, flies with the *for* variants *rover* or *sitter* differ in their locomotory behaviour on food, and in PKG levels⁵. In honeybees, the effect occurs through gene regulation rather than variation: when bees shift from working in the hive to foraging, *for* is upregulated⁶. Levels of PKG thus affect food-related behaviours over different timescales — over evolutionary time in *Drosophila* and over the lifetime of the individual honeybee.

So, has natural selection acted on different signalling molecules (the cGMP-gated ion channel in *C. elegans*, and PKG in flies and bees) in the same pathway to modulate food-related behaviours in these three species⁷? An obvious experiment is to determine whether *npr-1* and *for* lie in the same cGMP signalling pathway, and if they do whether there are similarities in the neural substrates involved. Whether cGMP signalling is involved in mammalian food-related behaviours also remains to be seen.

The discovery of an association between social behaviour and neurons that respond to aversive conditions is exciting because it provides a mechanistic basis for observations from behavioural ecology. A common theme in this rich literature is that adverse environmental conditions stimulate the formation of groups⁸. For example, solitary locusts become gregarious in response to drought, and birds of some species postpone mating to help their parents in the nest when new nest sites are scarce.

Many questions remain, of course. Is there a common neurobiological theme to group formation in response to stressful ecological conditions? Is the formation of social groups in general associated with neurons that sense aversive stimuli? Finally, if sociality has indeed evolved in response to stress, have the neurobiological underpinnings in simple organisms such as *C. elegans* provided the evolutionary basis for more complex social behaviours?

Marla B. Sokolowski is in the Department of Zoology, University of Toronto, 3359 Mississauga Road, Mississauga, Ontario L5L 1C6, Canada. e-mail: msokolow@utm.utoronto.ca

1. de Bono, M., Tobin, D. M., Davis, M. W., Avery, L. & Bargmann, C. I. *Nature* **419**, 899–903 (2002).
2. Coates, J. C. & de Bono, M. *Nature* **419**, 925–929 (2002).
3. Rankin, C. H. *Nature Rev. Genet.* **3**, 622–630 (2002).
4. de Bono, M. & Bargmann, C. I. *Cell* **94**, 679–689 (1998).
5. Osborne, K. et al. *Science* **277**, 834–836 (1997).
6. Ben-Shahar, Y., Robichon, A., Sokolowski, M. B. & Robinson, G. *Science* **296**, 741–744 (2002).
7. Sokolowski, M. B. *Neuron* **21**, 1–4 (1998).
8. Wilson, E. O. *Sociobiology: The New Synthesis* (Harvard Univ. Press, Cambridge, Massachusetts, 1975).



100 YEARS AGO

The Huxley Memorial Tablet... was unveiled at the Ealing Public Library on Thursday last by the Mayor of Ealing, Alderman H. C. Green. The inscription on the tablet is, “The Right Honourable Thomas Henry Huxley. Born at Ealing, 4th May, 1825. Died at Eastbourne, 29th June, 1895. Try to learn something about everything, and everything about something.”

ALSO...

The third annual Huxley memorial lecture was delivered at the Anthropological Institute on October 21 by Prof. D. J. Cunningham, F.R.S. The subject was “Right-handedness and Left-brainedness,” and Prof. Cunningham referred to the general interest which it presents to all students of anthropology. So far as available evidence goes, it seems probable that right-handedness was a characteristic of man at a very early period in his evolution... Investigation shows that right-handedness is due to a transmitted functional pre-eminence of the left brain... The greater part, if not the whole, of the motor incitations which lead to articulate speech go out from the speech centre which resides in the left cerebral hemisphere. In left-handed people, the predominance of the right cerebral hemisphere is accentuated by the transference to it of the active speech centre. Left-handed people, therefore, speak from the right brain. From *Nature* 30 October 1902.

50 YEARS AGO

The practice at the Royal Observatory, Greenwich, of dropping the Time Ball, the symbol of Greenwich time throughout the world, was resumed at 1 p.m. on October 26. The Time Ball is a dull red sphere, 5 ft. in diameter; it is part of the Wren Observatory building and was erected for the purpose of enabling masters of vessels proceeding down the Thames to adjust their chronometers. The Time Ball is raised half-way up a mast, on which it slides, at five minutes to one, hauled to the top at two minutes to one, and released electrically on the hour. It was the earliest means of making Greenwich mean time known to the public, but the system has now been superseded for all practical purposes by the six pips broadcast by the B.B.C. and by the speaking clock operated by the G.P.O. The Ball was operated continuously from 1833, when first installed, until the Observatory was evacuated during the Second World War. From *Nature* 1 November 1952.

Feminization of male frogs in the wild

Water-borne herbicide threatens amphibian populations in parts of the United States.

Atrazine is the most commonly used herbicide in the United States and probably in the world¹. Here we investigate the effects of exposure to water-borne atrazine contamination on wild leopard frogs (*Rana pipiens*) in different regions of the United States and find that 10–92% of males show gonadal abnormalities such as retarded development and hermaphroditism. These results are supported by laboratory observations, which together highlight concerns over the biological effects of environmental atrazine on amphibians.

We exposed *R. pipiens* larvae to different concentrations of atrazine (0, 0.1 or 25 parts per billion, p.p.b.) in the laboratory by immersion (30 larvae per treatment; $n = 3$) from just after hatching until tail resorption was complete. Only exposed males developed testicular oocytes (29% and 8%, respectively, at 0.1 and 25 p.p.b.); retarded gonadal development (gonadal dysgenesis) was evident in 36% and 12% of exposed males, respectively, and in one control animal (results not shown).

These findings are consistent with the more marked effects reported for endocrine-disruptors at lower doses (see ref. 2, for example). They also support previous indications that atrazine can cause gonadal abnormalities in males of *Xenopus laevis*^{3,4} and *Acris crepitans*⁵ in the laboratory. As its effects are not restricted to a single species, it is possible that this herbicide may pose a threat to amphibians in general.

We also examined leopard frogs, sampled from eight different sites in a transect running from Utah to Iowa, for abnormalities comparable to those seen under laboratory conditions. We used records of atrazine sales to identify potentially contaminated

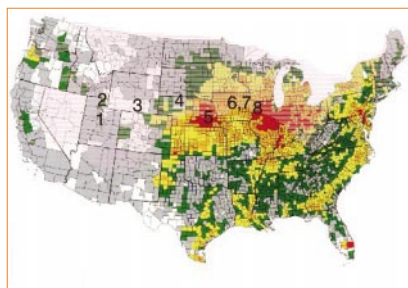


Figure 1 Use of the herbicide atrazine in the United States, on the basis of sales¹¹. Pink overlay shows the natural range of leopard frogs (*Rana pipiens*)^{12,13}. Regional coloration indicates atrazine usage (in kg km^{-2}): white, zero or no data; grey, <0.4 ; olive, $0.4\text{--}2.4$; yellow, $2.5\text{--}9.2$; orange, $9.3\text{--}28.7$; red, >28.7 . Numbers indicate the eight sites at which frogs were collected (for histological analysis of gonads) and water was sampled for atrazine analysis. We collected 100 animals at each site, selecting small individuals in order to sample newly metamorphosed animals, and killed them immediately in benzocaine; after fixation (Bouin's) for 48 h and preservation in 70% ethanol, animals were measured and their sex was determined. The histology of the gonads of 20 males and a subset of females taken from each site was analysed. Further details are available from the authors.

sites (Fig. 1). As control sites, we used various non-agricultural regions in Utah, Wisconsin and Nebraska that reported atrazine sales of less than 0.4 kg km^{-2} , as well as a non-agricultural area in Iowa. A golf-course pond in Cache county, Utah (the only county reporting atrazine sales of more than 0.4 kg km^{-2}), and cornfields in Nebraska and Iowa were considered to be likely sites of contamination. Water sampling revealed that only one site (Juab county, Utah) had atrazine levels below our detection limit (0.1 p.p.b.).

This site was the only locality where testicular oocytes were not observed in the

local population of leopard frogs. All sites associated with atrazine sales exceeding 0.4 kg km^{-2} and with water-borne atrazine contamination above 0.2 p.p.b. were found to contain males with testicular oocytes (Fig. 2a, b). These abnormalities were of similar morphology to those induced by atrazine in the same species in the laboratory. This hermaphroditism was not evident in the absence of atrazine exposure. We conclude that atrazine is responsible for these effects in wild populations, even though other contaminants may be present that could produce similar effects.

Atrazine may affect sex differentiation by inducing aromatase, the enzyme that converts androgens into oestrogens, and can cause inappropriate synthesis and secretion of oestrogens in males at the expense of androgens. This occurs in fish⁶, reptiles⁷ and mammals^{6,8}, with inhibition of spermatogenesis probably being a secondary effect associated with the depletion of androgens and synthesis of oestrogens in exposed males, rather than a direct effect of atrazine. Evidence for this mechanism of toxicity in three out of five vertebrate classes, and possibly in amphibians as well, generalizes the possible environmental risk associated with atrazine.

Most water sources in the United States, including rain, contain more atrazine than the effective doses determined in laboratory studies¹. Although the locality in Wyoming (North Platte River) with the highest frequency of sex reversal (92% of males) is not in the vicinity of farms and is not in a county that reports significant atrazine usage, hermaphrodite frogs are prevalent there because the North Platte River is fed by atrazine-contaminated⁹ streams that originate in Colorado.

The frequency of abnormalities at site 2 is much lower than at site 3, although the contamination measured at these sites was comparable. It may be that intermittently exposed populations are more susceptible to atrazine-induced hermaphroditism, whereas continuously exposed populations undergo adaptive resistance.

Applied to crop fields as a pre-emergent, atrazine contamination in water sources peaks with spring rains, which also coincide with breeding activity in many amphibians. Given the adverse effects of atrazine on the gonads of male frogs, this pattern of atrazine application may increase its impact on amphibian populations. In the light of growing evidence that these populations are in decline¹⁰, the contribution of atrazine to this decline warrants further investigation.

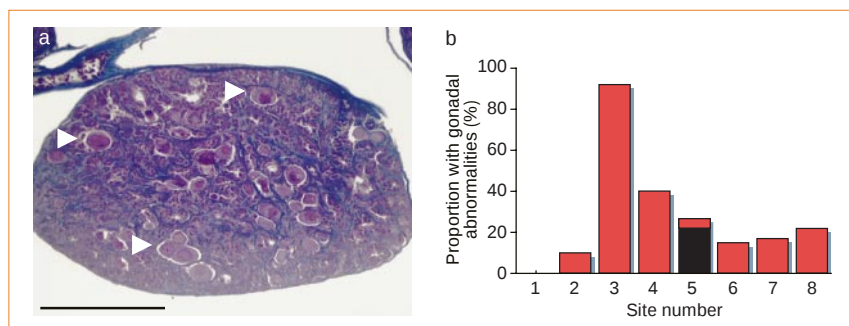


Figure 2 Testicular oogenesis in wild leopard frogs. **a**, Transverse histological cross-section ($8 \mu\text{m}$ thick; stained with Mallory's trichrome stain) of the left gonad of a male frog collected from Carbon County, Wyoming (site 3), showing testicular oocytes (arrowheads). More than 30 oocytes are present in this section. Serial sections reveal that every lobule has at least one oocyte, and some lobules (as shown) have as many as three. Scale bar, $250 \mu\text{m}$. **b**, Frequency of occurrence of gonadal dysgenesis (black; poorly developed testes that lack distinct lobules) and testicular oogenesis (hermaphroditism; red) in frogs collected from the eight different sites shown in Fig. 1. The water-borne contamination by atrazine (measured in p.p.b.) was: site 1, 0.14 ; site 2, 0.20 ; site 3, 0.20 ; site 4, 0.30 ; site 5, 0.80 ; site 6, 6.70 ; site 7, data not available; site 8, 0.5 .

Tyrone Hayes, Kelly Haston, Mable Tsui, Anhthu Hoang, Cathryn Haeffele, Aaron Vonk

Laboratory for Integrative Studies in Amphibian Biology, Museum of Vertebrate Zoology, and Department of Integrative Biology, University of California, Berkeley, California 94720-3140, USA e-mail: tyrone@socrates.berkeley.edu

1. US Environmental Protection Agency. *Federal Register* **59**, 60412–60443 (1994).
2. Akingbemi, B. T. & Hardy, M. P. *Ann. Med.* **33**, 391–403 (2001).
3. Hayes, T. B. *et al. Proc. Natl Acad. Sci. USA* **99**, 5476–5480 (2002).
4. Tevera-Mendoza, L. *et al. Environ. Toxicol. Chem.* **21**, 527–531 (2002).
5. Reeder, A. L. *et al. Environ. Health Perspect.* **106**, 261–266 (1998).
6. Sanderson, J. T., Letcher, R. J., Heneweer, M., Giesy, J. P. & van

- den Berg, M. *Environ. Health Perspect.* **109**, 1027–1031 (2001).
7. Crain, D. A., Guillette, L. J. Jr, Rooney, A. A. & Pickford, D. B. *Environ. Health Perspect.* **105**, 528–533 (1997).
8. Sanderson, J. T., Seinen, W., Giesy, J. P. & van den Berg, M. *Toxicol. Sci.* **54**, 121–127 (2000).
9. Kimbrough, R. A. & Litke, D. W. *Pesticides in Surface Water in Agricultural and Urban Areas of the South Platte River Basin from Denver, Colorado to North Platte, Nebraska, 1993–94* (NAWQA Program, South Platte River Basin Study, Denver, Colorado, 1995).
10. Wake, D. B. *Science* **253**, 860 (1991).
11. Battaglin, W. A. & Goolsby, D. A. *Water-Resources Invest. Rep.* 94–4176 (US Geol. Surv., Denver, Colorado, 1995).
12. Stebbins, R. A. *Field Guide to Western Reptiles and Amphibians: Field Marks of All Species in Western North America, Including Baja California* (Houghton–Mifflin, Boston, 1985).
13. Conant, R. A. *Field Guide to Reptiles and Amphibians: Eastern and Central North America* (Houghton–Mifflin, Boston, 1998).

Competing financial interests: declared none.

Brain development

Memory enhancement in early childhood

Regions of the brain's frontal lobe that are associated with memory retention and retrieval^{1,2} begin to mature during the last quarter of the first year in humans. This implies that infants younger than 8 or 9 months should have difficulty in registering an experience and retrieving it after a long delay^{3,4}. Here we show that 13-month-old children are unable to recall a sequence of actions performed in front of them when they were 9 months old, whereas 21- and 28-month-olds are able to retrieve representations of the same acts when these were witnessed at 17 and 24 months. Our findings indicate that long-term retention increases during the second year and support the idea that maturation of the frontal lobe at the end of the first year contributes to memory enhancement during this period.

Infants of 6 months old can remember events for up to 24 hours⁵, which extends to up to a month when they are 9 months old⁶. It has been proposed that early deficiencies in registering and retaining memories for events in the long term might be related to the protracted maturation of the neocortex⁷. In humans, the brain undergoes important changes towards the end of the first year, including the growth and differentiation of cortical pyramidal neurons and of dendrites in the hippocampus^{8–10}, which continue into the second year^{11,12}. These developmental processes should increase the efficiency of integration and registration of information in the neocortex, and in the prefrontal cortex in particular¹².

To test this hypothesis, we evaluated the ability of infants to retrieve representations after a delay of 4 months of motor acts first experienced at 9, 17 or 24 months of age. We used a deferred-imitation procedure in which the experimenter performed a sequence of actions with the aid of props while describing these actions verbally (for

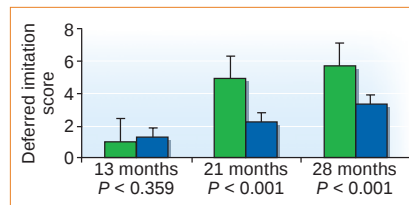


Figure 1 Deferred imitation by infants in three different age groups ($n=12$) for sets of familiar and new action-sequences. The mean deferred-imitation scores (sum of target actions and ordered pairs reproduced) are shown according to age group for familiar (left bars) and novel (right bars) sequences. Older infants perform better than younger ones on sequences that are new to them, presumably because they are better able to deduce the appropriate target actions without the aid of memory. Note that the performance of the 13-month-old group is weak for both familiar and new action-sequences. The 21- and 28-month-old groups, however, perform significantly better on sequences that they have seen 4 months earlier than they do on novel sequences. Further experimental details are available from the authors.

example, “Clean-up time!” was used for wiping the table with a paper towel and then throwing the towel into the waste basket; “Make a rattle!” was used for inserting a ring into a slot in a bottle and then shaking the bottle). We estimated the children’s recall of these events four months later on the basis of the number of actions they successfully re-enacted and on the number of pairs of actions performed in the proper sequence.

We exposed infants at 9, 17 or 24 months of age (12 in each group) to three of five possible action-sequences and encouraged them immediately to imitate each sequence. Infants aged 17 and 24 months witnessed four demonstrations of each sequence; 9-month-olds witnessed an additional two demonstrations (for a total of six demonstrations per sequence) to compensate for their immaturity.

After a 4-month delay, we tested the children’s ability to re-enact each sequence in response to the same verbal cues when presented with the props for all five action-sequences (the sequences were not demonstrated again after the delay). A comparison

of the children’s performance on the familiar and novel sequences served as a control for effects unrelated to memory. Five children from the first session were unable to participate in phase two, so their data were not included in the analysis. Six age-matched children, for whom all five action-sequences were novel in session two, were recruited to serve exclusively as controls.

As expected, the 21- and 28-month-olds showed a robust memory for events experienced 4 months earlier, whereas the 13-month-olds did not (Fig. 1). Subjects from the 21- and 28-month-old groups produced more target actions ($F(1,183) = 9.95$, $P < 0.002$) and more ordered pairs ($F(1,183) = 24.55$, $P < 0.001$) on familiar action-sequences than on novel ones. The 13-month-olds failed to produce a greater number of target actions ($t = -1.05$, $P < 0.15$) or correctly ordered pairs ($t = -0.27$, $P < 0.39$) on familiar sequences. In contrast, the 21-month-old ($t = 3.00$, $P < 0.002$) and 28-month-old ($t = 3.17$, $P < 0.001$) groups produced a significantly greater number of target actions on familiar sequences than on novel ones. This trend was even more pronounced for ordered pair scores, where 21-month-olds ($t = 4.60$, $P < 0.001$) and 28-month-olds ($t = 3.55$, $P < 0.001$) had consistently higher scores on sequences they had watched 4 months earlier.

Our findings that long-term memory in infants improves during their second year could be due to compromised registration or poor retrieval in the first year. In either case, our results support the popular belief that at 9 months the hippocampus and regions of the frontal cortex are not yet fully mature. They also indicate that there is a neurobiological component to memory enhancement across the second year, contrary to early assumptions that this is entirely attributable to experience¹³.

Conor Liston, Jerome Kagan

Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA e-mail: cliston@post.harvard.edu

1. Buckner, R. L. *et al. J. Neurosci.* **15**, 12–29 (1995).
2. Nyberg, L. *et al. Proc. Natl Acad. Sci. USA* **93**, 11280–11285 (1996).
3. Herschkowitz, N., Kagan, J. & Zilles, K. *Neuropediatrics* **28**, 296–306 (1997).
4. Herschkowitz, N., Kagan, J. & Zilles, K. *Neuropediatrics* **30**, 221–230 (1999).
5. Barr, R., Dowden, A., & Hayne, H. *Inf. Behav. Dev.* **19**, 159–170 (1996).
6. Carver, L. J. & Bauer, P. J. *J. Exp. Psychol.* **130**, 726–745 (2001).
7. Bachevalier, J., Brickson, M. & Hagger, C. *NeuroReport* **4**, 77–80 (1993).
8. Koenderink, M. J., Uylings, L. & Mrzljak, L. *Brain Res.* **653**, 173–182 (1994).
9. Mrzljak, L., Uylings, H. B., Van Eden, C. G. & Judas, M. *Prog. Brain Res.* **85**, 185–222 (1990).
10. Seress, L. & Mrzljak, L. *Hippocampus* **2**, 127–141 (1992).
11. Huttenlocher, P. R. *J. Comp. Neurol.* **387**, 167–178 (1997).
12. Van Buren, J. M. & Borke, R. C. *Variations and Connections of the Human Thalamus* (Springer, Berlin, 1972).
13. Piaget, J. *The Origins of Intelligence in Children* (transl. Cook, M.) (International Univ. Press, New York, 1952).

Competing financial interests: declared none.

Chemical mimicry

Male ants disguised by the queen's bouquet

Males of the tropical ant *Cardiocondyla obscurior* are either wingless and aggressive or winged and docile, and both compete for access to virgin queens in the nest^{1,2}. Although the fighter males (ergatoids) attack and kill other ergatoids, they tolerate and even attempt to mate with their winged rivals. Here we show that the winged males avoid the aggression of wingless males by mimicking the chemical bouquet of virgin queens, but that their mating success is not reduced as a result. This example of female mimicry by vigorous males is surprising, as in other species it is typically used as a protective strategy by weaker

males, and may explain the coexistence and equal mating success of two male morphs.

Ants typically mate during short nuptial flights that give little opportunity for male–male combat³. By contrast, sexual individuals of *C. obscurior* mate inside their natal nests. Intensive intrasexual competition has led to the evolution of sabre-shaped mandibles and escalated fighting to the death among locally mating ergatoid males^{1,2}. Winged males, on the other hand, have weak mandibles, behave peacefully and leave the nest one to two weeks after emergence in order to mate outside. But, surprisingly, before dispersal they are as successful as ergatoid males in attaining copulations with virgin nestmate queens (*U*-test: $U=903.0$, $P=0.86$, $n_{\text{ergatoid}}=33$ observations for 24 min each, $n_{\text{winged}}=56$). Ergatoid males would therefore be expected to benefit from killing their winged rivals⁴.

Contrary to the aggression that might be expected towards winged males, ergatoid males typically restrict their fighting to other ergatoids, and attempt to mate with the winged males¹. Our behavioural observations indicate that young winged males (1–5 days old) are as attractive as young virgin queens to ergatoid males (ergatoid males mounted 11 out of 12 winged males and 15 out of 17 virgin queens; 72–120 min observation per individual; Fisher's exact test: d.f. = 1, $P=1.0$). By contrast, ergatoid males showed no interest in older winged males (6–10 days, $n=17$; old–young: $P<0.001$), whereas virgin queens remained attractive regardless of their age (old: 12 out of 14; old–young: $P=1.0$).

This toleration of winged males and attempts at homosexual mating with them can be explained by the chemical resemblance of winged males and virgin queens in their bouquet of cuticular hydrocarbons on the body surface, which are important for communication in social insects^{5–8}. In a discriminant analysis of hydrocarbon profiles obtained by gas chromatography with mass spectrometry⁹, 1-day-old winged males and virgin queens formed a single cluster that was distinct from those of ergatoid males and workers (Fig. 1a). However, for 10-day-old ants, all four groups were clearly separated (Fig. 1b). We conclude that the odour similarity of young — and only young — winged males to virgin queens explains their age-dependent attractiveness to ergatoid males.

Most examples of chemical deception occur between species⁵, whereas intra-specific female mimicry is typically based on morphological or behavioural similarities¹⁰. Chemical female mimicry occurs in male rove beetles¹¹ and garter snakes¹² when they find themselves in adverse physiological conditions. These males pay for avoiding competition with 'high-quality' males by suffering low mating success with females.

By contrast, in *Cardiocondyla*, the odour mimicry of virgin queens by young winged males does not affect their copulatory success when compared to older, non-mimicking winged males ($U=324.0$, $P=0.30$, $n_{\text{young}}=25$, $n_{\text{old}}=31$) and ergatoid males ($U=393.0$, $P=0.76$).

Cardiocondyla provides a new evolutionary context for chemical deception: it is deployed irrespective of condition by all young winged males and renders them simultaneously attractive to both sexes. This example of chemical female mimicry enables two alternative reproductive strategies to exist in the male sex.

Sylvia Cremer*,†, Matthew F. Sledge‡, Jürgen Heinze*

*Biology 1, University of Regensburg, 93040 Regensburg, Germany

e-mail: sylvia.cremer@biologie.uni-regensburg.de

†Present address: Department of Population

Ecology, Zoological Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark

‡Department of Animal Biology and Genetics, University of Florence, 50125 Florence, Italy

1. Kinomura, K. & Yamauchi, K. *J. Ethol.* **5**, 75–81 (1987).
2. Heinze, J., Hölldobler, B. & Yamauchi, K. *Behav. Ecol. Sociobiol.* **42**, 239–246 (1998).
3. Hölldobler, B. & Bartz, S. H. in *Experimental Behavioral Ecology and Sociobiology* (eds Hölldobler, B. & Lindauer, M.) 167–220 (Academic, New York, 1985).
4. Anderson, C., Cremer, S. & Heinze, J. *Behav. Ecol.* (in the press).
5. Dettner, K. & Liepert, C. *Annu. Rev. Entomol.* **39**, 129–154 (1994).
6. Lahav, S., Soroker, V. & Hefetz, A. *Naturwissenschaften* **86**, 246–249 (1999).
7. Lenoir, A., D'Ettorre, P., Errard, C. & Hefetz, A. *Annu. Rev. Entomol.* **46**, 573–599 (2001).
8. Singer, T. L. *Am. Zool.* **38**, 394–405 (1998).
9. Turillazzi, S., Sledge, M. F., Cremer, S. & Heinze, J. *Insect Soc. Life* **4**, 169–175 (2002).
10. Andersson, M. in *Sexual Selection* (eds Krebs, J. R. & Clutton-Brock, T.) Ch. 16 (Princeton University Press, Princeton, New Jersey, 1994).
11. Peschke, K. *J. Chem. Ecol.* **13**, 1993–2008 (1987).
12. Shine, R., Harlow, P., LeMaster, M. P., Moore, I. T. & Mason, R. T. *Anim. Behav.* **59**, 349–359 (2000).
13. Heinze, J., Stengl, B. & Sledge, M. F. *Behav. Ecol. Sociobiol.* **52**, 59–65 (2002).

Competing financial interests: declared none.

COMMUNICATIONS ARISING

Palaeoclimatology

Tropical temperatures in greenhouse episodes

Reconstruction from oxygen-isotope data of global sea-surface temperatures (SSTs) that prevailed during periods when greenhouse-gas levels were higher than at present (that is, in the early Eocene and Cretaceous epochs) indicates that temperatures were significantly warmer than present in mid- to high latitudes, but similar to or slightly cooler in the tropics. However, Pearson and colleagues¹ claim, on the basis of oxygen-isotope data from planktonic foraminifers recovered from clay beds in Tanzania and other localities, that the tropics were warm during

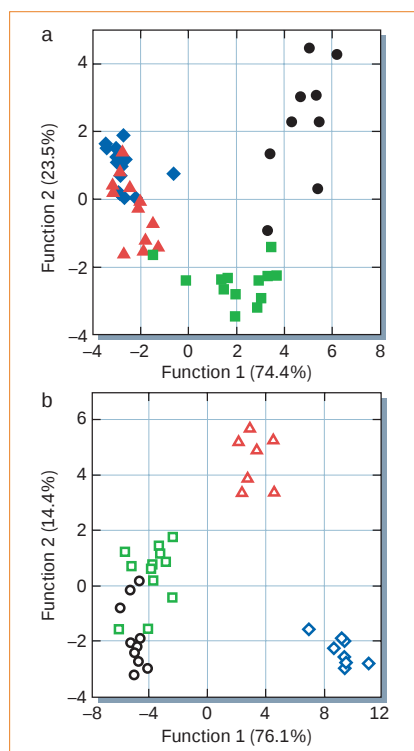


Figure 1 Discriminant analysis of the hydrocarbon profiles of 1-day-old (filled symbols) and 10-day-old (hollow symbols) *Cardiocondyla obscurior* ants. Hydrocarbons were extracted from workers (day 1: $n_1=10$, day 10: $n_{10}=10$; circles), ergatoid males ($n_1=13$, $n_{10}=12$; squares), winged males ($n_1=14$, $n_{10}=8$; diamonds) and virgin queens ($n_1=13$, $n_{10}=7$; triangles) by gas chromatography with mass spectrometry, and were statistically analysed by principal-components analysis followed by stepwise discriminant analysis¹³. Results show the first two canonical discriminant functions (percentage of variability is shown in parentheses). **a**, One-day-old ants: virgin queens and winged males cluster together but are distinct from ergatoid males and workers (81.6% of individuals classified correctly; F1: Wilks' $\lambda=0.018$, $P<0.001$; F2: Wilks' $\lambda=0.194$, $P<0.005$). **b**, Ten-day-old ants: all colony members form distinct groups (as the non-plotted third significant function discriminates workers and ergatoid males; 100% of individuals classified correctly; F1: Wilks' $\lambda=0.001$, $P<0.001$; F2: Wilks' $\lambda=0.023$, $P<0.001$).

the Eocene, attributing the discrepancy between these and previous tropical SST estimates to the effects of diagenetic overprinting of the latter. Here we note some potential flaws in their interpretation of the new oxygen-isotope data.

Although dissolution and incipient secondary calcification of tropical, deep-sea fossil planktonic foraminifera bias shell $\delta^{18}\text{O}$ ratios towards higher values^{2–4}, we consider it unlikely that all tropical planktonic foraminifera tests are overprinted to the extent claimed by Pearson *et al.* Their estimate of up to 50% secondary calcite is based on a comparison of multi-species carbon and oxygen isotope values from planktonic foraminifera of “similar age” from Tanzanian outcrops and the Deep Sea Drilling Project Site 523 in the eastern South Atlantic Ocean, and their key assumption is that initial conditions are identical for both sets of samples.

However, given the differences in latitude, and the likely proximity of Tanzania and Site 523 to warm western and to cool eastern boundary currents, respectively, SST could have differed by as much as 6–8 °C between the two localities in the Eocene, just as it does today. This alone could account for 80% of the $\delta^{18}\text{O}$ difference between the two data sets. Moreover, as implied by the clay-rich depositional facies from which the Tanzanian fossils were extracted, it is likely that regional sea-surface salinities and seawater $\delta^{18}\text{O}$ were lower, possibly by as much as 3.0 p.p.t. and 1.0 p.p.t., respectively, than at Site 523. This local salinity difference would further bias the Tanzanian SST estimates towards higher values. These biases are compounded by the fact that the chronostratigraphic constraints on the Tanzanian sediment sequences are relatively coarse, a point that is reflected in the reported age errors of the Tanzanian samples, which is $\pm 1\text{--}2$ Myr.

As a result of these and other potential biases, it may prove difficult to constrain absolute estimates of SST in tropical coastal oceans to better than ± 3.0 °C for Eocene or Cretaceous ‘greenhouse’ intervals. This raises the issue of whether this is the most effective approach to resolving the enigma of greenhouse-gas tropical SSTs. We think not, and favour a more practical strategy that focuses on relative changes in temperature that were associated with episodes of rapid, short-lived climate change^{5–7}.

Why? One reason is that it is easier to quantify relative changes in temperature on short timescales (about $10^2\text{--}10^4$ yr), even in sediments that have been slightly to moderately overprinted. Because secondary calcification in sediments of uniform lithology is constant over short length scales, primary offsets in fossil-shell chemistry tend to be preserved, although slightly attenuated^{3,6}.

Second, by focusing on rapid, transient

events, the number of climate-forcing parameters can be reduced to greenhouse-gas levels, as all other non-astronomical parameters, such as changes in continental geography and elevation, operate on much longer timescales. However, this strategy will be most effective when dealing with well-dated, stratigraphically continuous (over millions of years) and lithologically uniform sedimentary sequences, such as those typically recovered from the deep sea.

James C. Zachos*, Michael A. Arthur†, Timothy J. Bralower‡, Howard J. Spero§

*Department of Earth Sciences, University of California, Santa Cruz, California 95064, USA
e-mail: jzachos@earthsci.ucsc.edu

†Department of Geosciences, Pennsylvania State University, State College, Pennsylvania 16802, USA

‡Department of Geological Sciences, University of North Carolina, Chapel Hill, North Carolina 27599, USA

§Department of Geology, University of California, Davis, California 95616, USA

1. Pearson, P. N. *et al.* *Nature* **413**, 481–485 (2001).
2. Wu, G., Herguera, J. C. & Berger, W. H. *Paleoceanography* **5**, 581–594 (1990).
3. Schrag, D. P., dePaolo, D. J. & Richter, F. M. *Geochim. Cosmochim. Acta* **59**, 2265–2278 (1995).
4. Wilson, P. A. & Opdyke, B. N. *Geology* **25**, 555–558 (1996).
5. Kennett, J. P. & Stott, L. D. *Nature* **353**, 225–229 (1991).
6. Kelly, D. C., Bralower, T. J. & Zachos, J. C. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **141**, 139–161 (1998).
7. Thomas, D. J., Bralower, T. J. & Zachos, J. C. *Paleoceanography* **14**, 561–570 (1999).

Pearson *et al.* reply — We agree with Zachos and colleagues that much can be learned about transient climate events from well-dated carbonate sediments from the deep ocean, and that the problem of recrystallization of such sediments (when tens of millions of years old) means that constraining relative temperature changes from oxygen-isotope data is more appropriate than attempting to measure absolute values.

We note, however, that this marks a shift of emphasis from studies spanning several decades, in which attempts have been made to determine actual temperatures from the deep-sea record and to compare them with the output of climate-model simulations (for example, see refs 1–3). We also warn against assuming that “primary offsets in fossil shell chemistry tend to be preserved”, as diagenetic overprinting will tend markedly to reduce such offsets, as well as attenuating stratigraphic patterns.

We have suggested that recrystallization could account for as much as 50% of the mass of planktonic foraminifera in typical deep-sea samples, which would reduce estimated SSTs in the tropics by 10–15 °C. Zachos *et al.* have questioned our comparison of isotope data from well-preserved, middle-Eocene foraminifera from Tanzania and a recrystallized deep-sea assemblage. But we did not assume that initial conditions

were identical for the two sites — indeed, we were careful to point out the likely original difference in temperature. When this is taken into account, we calculate that the oxygen-isotope data suggest a 55% diagenetic overprint of the deep-sea sample towards an ocean bottom-water temperature of about 10 °C (ref. 4). A similar overprint is also implied by the greatly reduced interspecies carbon-isotope differentials.

We also discussed the question of $\delta^{18}\text{O}_{\text{sw}}$ on the Tanzanian margin. For several reasons (the narrow open shelf, probable onshore current, normal stratification, full plankton assemblages and results of palaeosalinity modelling), we think it unlikely that the salinity was as much as 3.0 p.p.t. lower than in the deep ocean. But even if it were, the probable relationship between $\delta^{18}\text{O}_{\text{sw}}$ and salinity at this latitude would imply a $\delta^{18}\text{O}_{\text{sw}}$ value that is only about 0.5 p.p.t. more negative, which corresponds to an estimated temperature increase of about 2 °C. This is a small effect when compared with the apparent difference of nearly 15 °C in estimated SST between the East African shelf and coeval open-ocean sites such as ODP Site 865 (ref. 2). We stress that our Tanzanian data are also supported by similar results from Mexico, Alabama and the Adriatic Sea.

Study of both deep-sea carbonates and hemipelagic clays is crucial to ensure that sampling is as spatially distributed as possible. But there is no reason why hemipelagic mudstones should not be as accurately dated as deep-sea carbonates, thereby combining the advantages of high-resolution stratigraphy and good microfossil preservation. Recent Ocean Drilling Program (ODP) coring in New Jersey has already demonstrated this potential⁵, and we anticipate further well-dated Cretaceous and Palaeogene mudstones from a forthcoming ODP leg to Demerara Rise (in the equatorial Atlantic Ocean) and from our own drilling in Tanzania. Such investigations will help to elucidate past variation in absolute palaeotemperatures and meridional temperature gradients, which remains critical for testing the greenhouse theory for past warm climates.

Paul N. Pearson*, Peter Ditchfield*, Nicholas J. Shackleton†

*Department of Earth Sciences, University of Bristol, Queen's Road, Bristol BS8 1RJ, UK
e-mail: paul.pearson@bristol.ac.uk

†Godwin Laboratory, Department of Earth Sciences, University of Cambridge, Cambridge CB2 3SA, UK

1. D'Hondt, S. & Arthur, M. A. *Science* **271**, 1838–1841 (1996).
2. Bralower, T. *et al.* *Paleoceanography* **10**, 841–865 (1995).
3. Crowley, T. J. & Zachos, J. C. in *Warm Climates in Earth History* (eds Huber, B. T., MacLeod, K. S. & Wing, S. C.) 50–76 (Cambridge Univ. Press, Cambridge, 2000).
4. Oberhänsli, H. & Toumarkine, M. in *South Atlantic Paleoclimatology* (eds Hsu, K. J. & Weissert, H. J.) 125–148 (Cambridge Univ. Press, Cambridge, 1985).
5. Miller, K. G. *et al.* *Proc. Ocean Drilling Prog., Init. Rep. vol. 174AX* (Ocean Drilling Prog., College Station, Texas, 1998).

Social feeding in *Caenorhabditis elegans* is induced by neurons that detect aversive stimuli

Mario de Bono^{*†}, David M. Tobin^{*}, M. Wayne Davis^{‡§}, Leon Avery[‡] & Cornelia I. Bargmann^{*}

^{*} Howard Hughes Medical Institute, Programs in Developmental Biology, Neuroscience, and Genetics, Departments of Anatomy and Biochemistry & Biophysics, UCSF, California 94143-0452, USA

[†] MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[‡] Department of Molecular Biology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, Texas 75390-9148, USA

Natural *Caenorhabditis elegans* isolates exhibit either social or solitary feeding on bacteria. We show here that social feeding is induced by nociceptive neurons that detect adverse or stressful conditions. Ablation of the nociceptive neurons ASH and ADL transforms social animals into solitary feeders. Social feeding is probably due to the sensation of noxious chemicals by ASH and ADL neurons; it requires the genes *ocr-2* and *osm-9*, which encode TRP-related transduction channels, and *odr-4* and *odr-8*, which are required to localize sensory chemoreceptors to cilia. Other sensory neurons may suppress social feeding, as social feeding in *ocr-2* and *odr-4* mutants is restored by mutations in *osm-3*, a gene required for the development of 26 ciliated sensory neurons. Our data suggest a model for regulation of social feeding by opposing sensory inputs: aversive inputs to nociceptive neurons promote social feeding, whereas antagonistic inputs from neurons that express *osm-3* inhibit aggregation.

In many animal species individuals live in groups for part or most of their lives^{1,2}. Environmental signals such as predators, food availability and season can regulate aggregation behaviour^{1,2}. In turn, aggregation can profoundly influence population dynamics, community structure and biodiversity^{3,4}. Although our understanding of the ecological and evolutionary significance of aggregation has advanced considerably, the neural mechanisms underlying this form of behaviour have rarely been explored.

Many nematode species can aggregate, both in culture and in the wild^{5,6}. Natural isolates of the soil nematode *C. elegans* feed on bacteria either alone or in groups^{7,8}. Solitary feeders such as the standard N2 strain reduce locomotory activity and disperse on encountering bacterial food. By contrast, social feeders such as strain CB4856 continue moving rapidly on food and aggregate together. Social feeders also accumulate on the border of a bacterial lawn⁸. These behavioural differences are associated with natural variation at a single amino acid residue of NPR-1, a putative neuropeptide receptor related to the neuropeptide Y (NPY) receptor family⁸. Solitary strains bear valine at residue 215 of the *npr-1* gene, whereas social strains bear phenylalanine. Animals lacking *npr-1* strongly aggregate, indicating that *npr-1* represses this behaviour. Because the laboratory standard N2 strain is solitary, little is known about the sensory signals or neural mechanisms that promote social behaviour in *C. elegans*. Thus, *npr-1* mutants provide an opportunity to investigate this behaviour.

Stimulation of aggregation behaviour

To ascertain conditions required for aggregation of *npr-1* animals, we varied cultivation conditions and population density. Social feeders aggregated within 20 min of being placed on a fresh food source and remained in groups, whereas animals from the solitary strain N2 did not aggregate, or showed weak transient aggregation and then dispersed (Fig. 1a). Increasing population density, which favours encounters with other individuals, stimulated aggregation in *npr-1* strains (Fig. 1a).

Well-fed animals from both social and solitary strains disperse when food is absent. To investigate whether social animals would aggregate in the absence of food if they were maintained at a high population density, we trapped animals using a chemical or physical fence (see Methods). In the absence of food neither well-fed *npr-1* (*ad609*) (where *ad609* is a loss-of-function allele) mutants nor well-fed wild social animals showed a significant tendency to aggregate at any density (Fig. 1b).

One possible mediator of aggregation is a pheromone that induces dauer larva formation. Production of dauer pheromone requires the *daf-22* gene⁹. Social feeding by *npr-1* animals was not suppressed by a mutation in *daf-22*, or by mutations in other dauer pathway genes that affect a transforming growth factor (TGF)- β signalling cascade¹⁰ (*daf-3*, -5, -7, -8) or an insulin receptor signalling pathway (*daf-2*, *daf-16*)⁹ (data not shown). Thus the aggregation signal is unlikely to correspond to the dauer pheromone.

Mutants in *daf-1*, *daf-7*, *daf-8* and *daf-14* aggregate and border on food, although less strongly than *npr-1* mutant animals¹¹. This observation prompted a further investigation of the dauer pathway's relationship to *npr-1*. The aggregation and bordering behaviour of these *daf* mutants is completely suppressed by mutations in *daf-3* and *daf-5* (ref. 11), whereas *daf-3* and *daf-5* mutations did not inhibit aggregation and bordering of *npr-1* mutant animals (data not shown). To investigate whether the *daf* genes and *npr-1* regulate social feeding by a common pathway, we compared social feeding in adult *npr-1* and *daf-7* mutant animals with the behaviour of adult *daf-7*; *npr-1* double mutant animals. *daf-7* encodes a TGF- β homologue that is expressed in the ASI neurons in well-fed animals, but is transcriptionally repressed at high nematode population density by dauer pheromone^{10,12}. Social behaviour was enhanced in the *daf-7*; *npr-1* double mutants (Fig. 1c–f): the average number of animals in a group is greatly increased in *daf-7*; *npr-1*(null) strains compared with *npr-1*(null) or mutant *daf-7* strains alone. This result suggests that *daf-7* and *npr-1* act in parallel pathways to repress social feeding. For example, *daf-7* may mediate some of the effect of density on aggregation.

§ Present address: Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, Utah 84112-0840, USA.

Mutations in OSM-9 and OCR-2 disrupt social feeding

The observation that aggregation behaviour is modulated by food and population density initiated a search for sensory mechanisms that might mediate these effects. *Caenorhabditis elegans* responds to mechanosensory stimuli, chemical attractants and repellants, and temperature gradients^{13,14}. We asked whether mutations that disrupt one or more of these behaviours prevent *npr-1* mutant animals from aggregating. Mutants with defective responses to water-soluble attractants (*che-13*, *osm-1*, *osm-3*), attractive odorants (*che-13*, *osm-1*, *odr-1*, *odr-2*, *odr-3*, *odr-7*), aversive chemicals (*che-13*, *osm-1*, *osm-3*, *odr-3*), gentle touch on the body (*mec-3*, *mec-4*), touch to the head (*che-13*, *osm-1*), mechanical properties of bacteria (*cat-4*), or temperature (*ttx-3*) were still able to aggregate efficiently (Fig. 2a, and data not shown). Bordering behaviour on a bacterial lawn was also unaffected by these mutations (Fig. 2b, and data not shown).

Mutations in the *osm-9* and *ocr-2* genes abolished aggregation of *npr-1* mutant animals, and substantially reduced accumulation at the border of a bacterial lawn (Fig. 2c, d). *osm-9* and *ocr-2* encode predicted TRPV cation channel subunits that are thought to form a sensory transduction channel in several *C. elegans* chemosensory neurons^{15,16}. *osm-9* and *ocr-2* are required for some forms of olfactory chemoattraction as well as avoidance of hyperosmotic stimuli, nose touch, and certain volatile aversive stimuli. The closest

vertebrate relatives of OSM-9 and OCR-2 are also sensory channels: the vanilloid (capsaicin) receptor TRPV1 (VR1) that is implicated in sensing thermal pain, and the osmo-mechanosensitive channel TRPV4 (refs 17, 18).

The *odr-4* and *odr-8* genes are required for some olfactory responses, and act to localize a subset of olfactory receptors to olfactory cilia¹⁹. *odr-4* encodes a transmembrane protein, whereas the molecular identity of *odr-8* is unknown. Mutations in both *odr-4* and *odr-8* strongly disrupted the ability of *npr-1* mutant animals to aggregate and to border on food (Fig. 2c, d).

Mutations in *ocr-2*, *osm-9*, *odr-4* and *odr-8* could disrupt aggregation of social *npr-1* animals by reducing production of the stimuli that promote aggregation, or by reducing the detection of these stimuli. To distinguish between these two possibilities we investigated whether individual *npr-1* animals defective in *ocr-2*, *osm-9*, *odr-4* or *odr-8* function would join groups of social *npr-1* animals. All four double mutants, like the solitary wild strain N2, failed to join groups of *npr-1* social animals (Fig. 2e). These results suggest that the double mutants cannot sense signals that promote aggregation. Bordering behaviour of the double-mutant strains was also unaffected by the presence of social *npr-1* animals (Fig. 2f).

Animals from wild solitary strains such as N2 respond to food by strongly reducing their locomotory activity. This slowing response is mediated by a dopamine-containing neural circuit that senses a mechanical attribute of bacteria, and is enhanced by food-deprivation through a mechanism that involves serotonin²⁰. *npr-1* mutants and wild social strains do not show this slowing response, even after food deprivation (Fig. 3)⁸; however, double mutants between *npr-1* and *ocr-2* or *osm-9*, displayed a slowing response to food similar to that of N2 solitary animals (Fig. 3). *ocr-2*; *npr-1* and *osm-9*; *npr-1* animals also displayed the enhanced slowing response to food after food deprivation. These results suggest that *npr-1* is not

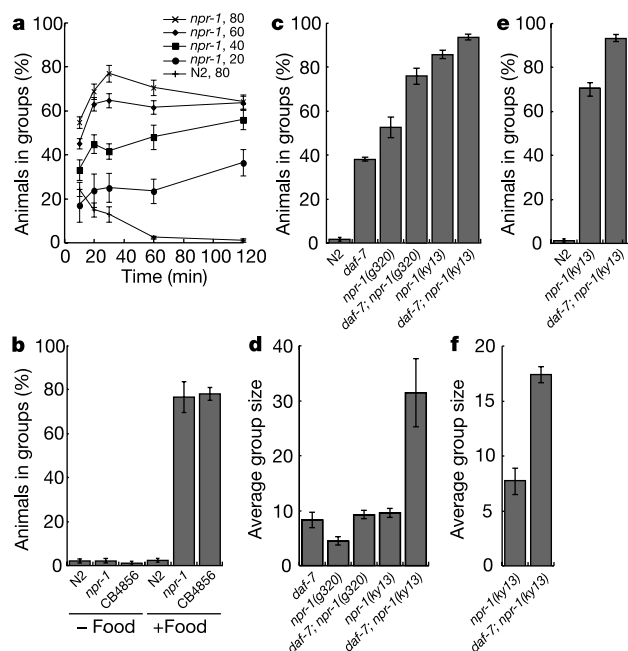


Figure 1 Aggregation of *npr-1* mutant animals requires food and is enhanced by increased population density and *daf-7* TGF- β mutations. **a**, Time courses for aggregation of 20, 40, 60 or 80 *npr-1(ad609)* animals, or of 80 N2 animals, as indicated. Populations of 20, 40 and 60 N2 animals show less aggregation than 80 N2 animals (data not shown). **b**, Well-fed social feeders do not aggregate without food. Population density in these assays was about 26 animals per cm². The *npr-1* allele is *ad609*; CB4856 is a wild social strain bearing the *npr-1* 215F allele. **c**, Aggregation behaviour with 80 adult animals per assay. *ky13* is a predicted null allele of *npr-1*; *g320* is the natural *npr-1* 215F allele. The *daf-7(e1372)* allele is a point mutation that acts like a strong loss-of-function allele¹⁰. *daf-7(e1372)* enhances the aggregation of both *npr-1(ky13)* ($P < 0.02$) and *npr-1(g320)* animals ($P < 0.01$). **d**, Average number of animals in a group for the assays in **c**. Only animals in groups were scored to get the average group size. *daf-7(e1372)*; *npr-1(ky13)* animals form significantly larger groups than *npr-1(ky13)* animals ($P < 0.05$). **e**, **f**, Per cent aggregation and average group size in assays involving a population of 40 adult animals. *daf-7(e1372)*; *npr-1(ky13)* animals aggregate more strongly and form larger groups than *npr-1(ky13)* animals ($P < 0.001$ for both behaviours). In all panels $n \geq 5$ assays.

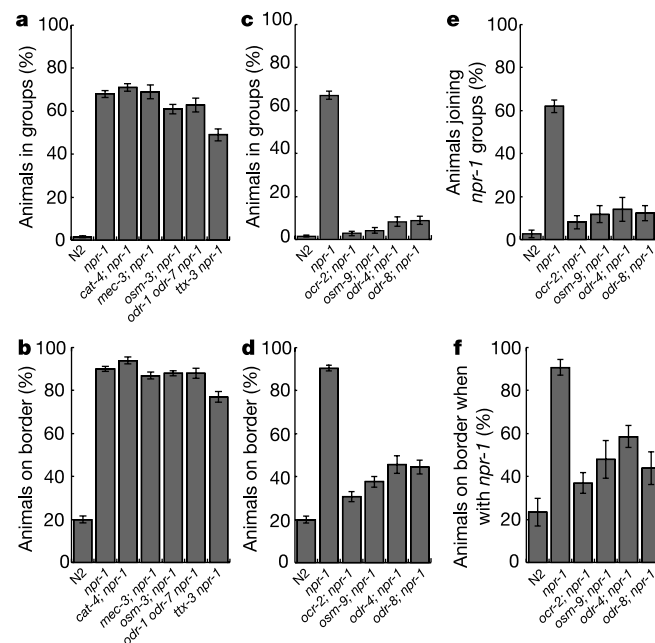


Figure 2 Social feeding of *npr-1* is lost in *ocr-2*, *osm-9*, *odr-4* or *odr-8* mutant backgrounds. **a**, **b**, *cat-4(e1141)*, *mec-3(e1338)*, *osm-3(p802)*, *odr-1(n1936)* *odr-7(ky4)* and *ttx-3(ks5)* mutations do not disrupt aggregation or bordering of *npr-1(ad609)* animals. The *odr-1 odr-7 npr-1* and *ttx-3 npr-1* animals also bear the *lon-2(e678)* X mutation. **c**, **d**, Mutations in *ocr-2*, *osm-9*, *odr-4* and *odr-8* strongly reduce aggregation and bordering of *npr-1* animals ($P < 0.001$). **e**, **f**, Individual animals of each genotype were tested for their ability to join groups of marked *npr-1* animals, or to border in the presence of *npr-1* animals (see Supplementary Information). For **a–d**, $n \geq 9$ population assays; for **e**, **f**, $n \geq 30$ individual animals, each tested 2–3 times.

essential for a food-induced slowing response. Instead, loss of *npr-1* may activate locomotion in a way that bypasses or suppresses its downregulation by food.

Various slow-moving *npr-1* double-mutant strains showed strong aggregation and bordering behaviours, including *mec-3(e1338); npr-1(ad609)*, *unc-25(e156); npr-1(ad609)*, *unc-30(e191); npr-1(ad609)* and *unc-36(e251); npr-1(ad609)*. Thus high locomotory activity is not necessary for social feeding behaviour.

OCR-2 and ODR-4 are required in nociceptive neurons

The expression patterns of the sensory transduction genes provided a set of candidate neurons for social feeding behaviour. OCR-2 and OSM-9 are co-expressed in six pairs of sensory neurons: ADF, ADL, ASH, AWA, PHA and PHB. To identify the cells in which OCR-2 activity is required for social behaviour, we expressed an *ocr-2* minigene in subsets of these six neurons using neuron-specific promoters. Expression of OCR-2 in either ASH or ADL neurons was able to restore social behaviour to *ocr-2; npr-1* animals; expression in the other types of neurons did not alter the solitary behaviour of *ocr-2; npr-1* animals (Fig. 4a).

The *odr-4* gene is expressed in 12 pairs of neurons, including ASH and ADL¹⁹. To investigate where *odr-4* activity is required for social behaviour, we expressed functional green fluorescent protein (GFP)-tagged ODR-4 in subsets of these 12 neurons in *odr-4; npr-1* animals, and assayed aggregation and bordering behaviour. Expression of *odr-4* in ADL, but not other neurons, restored both aggregation and bordering behaviours of *odr-4; npr-1* animals (Fig. 4b). This analysis of *ocr-2* and *odr-4* implicates the ASH and ADL neurons in social feeding behaviour.

In the *C. elegans* olfactory neuron AWA, the OCR-2/OSM-9 channel is thought to be gated by a G-protein-coupled cascade activated by odorant binding to olfactory receptors such as ODR-10 (refs 15, 21, 22). Similar pathways may operate in the ASH and ADL neurons, which also express predicted chemosensory receptors²³. ODR-4 acts to localize a subset of chemosensory receptors, including ODR-10, to the cilia of olfactory neurons¹⁹. The requirement of *odr-4* in ADL neurons for social feeding suggests that chemosensory receptors act in ADL to support this behaviour.

Ablation of ASH and ADL abolishes social feeding behaviour

To investigate whether the ASH and ADL neurons are required for social feeding, we ablated these neurons using a laser microbeam. The aggregation behaviour of ablated animals expressing a GFP marker was examined after mixing them with unmarked *npr-1* mutant animals. Ablation of both ASH neurons, or of both ADL neurons, did not disrupt social behaviour. However, simultaneous ablation of all ASH and ADL neurons strongly disrupted the ability

of *npr-1* animals to aggregate (Fig. 4c). These data suggest that input from either the ASH neurons or the ADL neurons is required for social feeding behaviour.

The ASH and ADL neurons are implicated in aversive responses to noxious stimuli. ASH neurons are important for avoidance of touch to the head, hyperosmotic solutions, extracts of dead worms, high concentrations of benzaldehyde, and toxic heavy metal ions such as Cu^{2+} and Cd^{2+} (refs 23–26). The ADL neurons are important for long-range avoidance of 1-octanol and contribute with ASH to short-range avoidance of 1-octanol^{24,27}. The well-documented roles for ASH and ADL in mediating responses to aversive stimuli suggests that such stimuli have a function in social behaviour: ASH and ADL may activate social feeding behaviour in response to repulsive cues. The most probable source of the aversive cues in standard culture conditions is bacteria.

To investigate whether bacterial signals induce social feeding, we assayed *npr-1* mutant animals on a lawn of *Escherichia coli* that had been killed using streptomycin. *npr-1* mutant animals became solitary feeders under these conditions (data not shown). Notably, *npr-1* mutant animals feeding on bacteria killed by streptomycin became social feeders when they were exposed to the odour of live bacteria placed on the lid of the assay plate. N2 animals continued to feed alone under these conditions. These results suggest that a bacterial odour induces social feeding behaviour.

Nociceptive neurons are antagonized by other neurons

Ablation of the ASH and ADL pairs of nociceptive neurons strongly disrupted social feeding of *npr-1* animals; however, a mutation in *osm-3* kinesin that truncates the sensory cilia of ASH and ADL²⁸ had no effect on social behaviour (Figs 2a, b and 5a, b). *osm-3* affects development of the distal cilia in 26 *C. elegans* sensory neurons that are open to the external environment²⁹, including ASH and ADL. As *osm-3* acts in other neurons apart from ASH and ADL, we reasoned that high levels of social behaviour in *osm-3; npr-1* animals could reflect a balanced removal of positively and negatively acting signals that regulate aggregation (Fig. 5c). In this model disruption of ASH and ADL neurons only removes stimuli that activate social behaviour, whereas *osm-3* mutations remove both the ADL/ASH positive signals and opposing negative ones (Fig. 5c). This hypothesis predicts that the triple mutants *odr-4; osm-3; npr-1* and *osm-3 ocr-2; npr-1* should aggregate indistinguishably from *osm-3; npr-1*

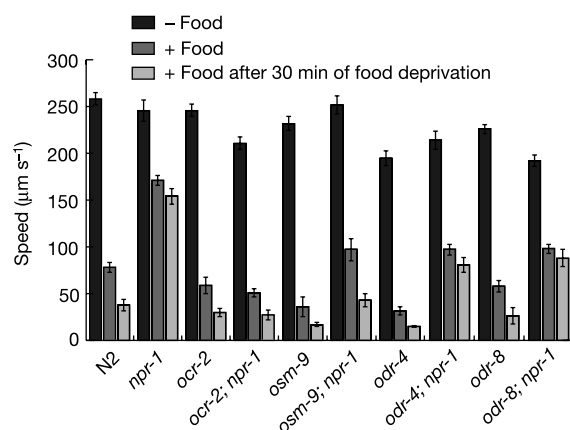


Figure 3 Mutations in *ocr-2*, *osm-9*, *odr-4* and *odr-8* restore the ability of *npr-1* mutant animals to slow down upon encountering food. $n \geq 25$ animals each recorded for 4 min.

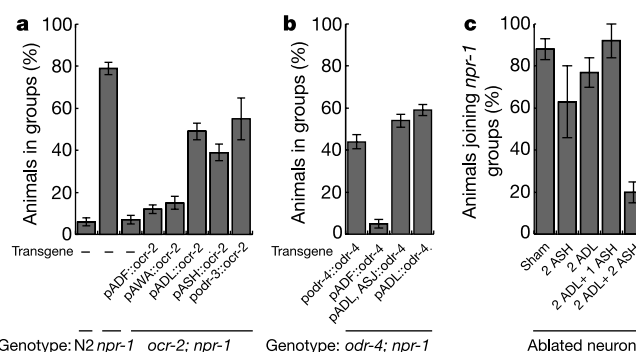


Figure 4 The nociceptive neurons ASH and ADL are required for social feeding behaviour. **a**, Expression of *ocr-2* in the ASH or ADL neurons restores aggregation behaviour to *ocr-2; npr-1* mutant animals ($P < 0.003$ compared to *ocr-2; npr-1*), whereas expression in the ADF or AWA neurons does not. *odr-3* is expressed in ASH, AWA, AWB, AWC and ADF neurons. **b**, *odr-4* expression in ADL restores social behaviour to *odr-4; npr-1* mutant animals ($P < 0.001$ compared to *odr-4; npr-1*), whereas expression in ADF does not. Expression of *odr-4* in AWA, AWB, ASL and ASK neurons also fails to restore aggregation behaviour to *odr-4; npr-1* animals (data not shown). **c**, Ablation of both pairs of ADL and ASH neurons suppresses aggregation ($P < 0.001$ compared to sham). Ablation of the ASH neuron pair, the ADL neuron pair, or the ASH neuron pair and 1 ADL neuron, does not suppress aggregation.

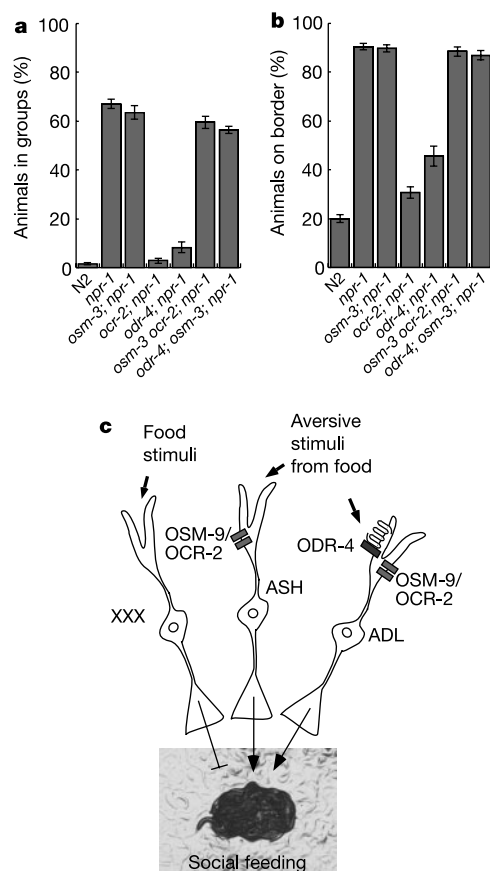


Figure 5 Disruption of *osm-3* kinesin restores social feeding to mutants defective in nociceptive neuron function. **a**, **b**, Aggregation (**a**) and bordering (**b**) behaviours were scored in animals of the indicated genotype. *osm-3 ocr-2; npr-1* and *odr-4; osm-3; npr-1* animals aggregate and border strongly to similar levels as *npr-1* and *osm-3; npr-1* animals ($P > 0.05$). $n \geq 10$ assays. **c**, A model for social feeding in *C. elegans*. The ASH and ADL nociceptive neurons are proposed to respond to aversive stimuli from food to promote social feeding. This function requires the putative OCR-2/OSM-9 ion channel. The ODR-4 protein may act in ADL to localize seven transmembrane domain chemoreceptors that respond to noxious stimuli. In the absence of ASH and ADL activity, an unidentified neuron (XXX) represses social feeding, perhaps in response to a different set of food stimuli. The photograph shows social feeding of a group of >30 *npr-1* mutant animals on a lawn of *E. coli*.

animals. This prediction is borne out: the triple mutant strains aggregated and bordered to the same level as *npr-1* and *osm-3; npr-1* mutants (Fig. 5a, b). Moreover, both *osm-3* and *osm-3; ocr-2* animals were solitary feeders if they expressed *npr-1* bearing valine at residue 215 (data not shown), indicating that the suppression of social feeding by the *npr-1* gene does not require *osm-3*.

These data suggest that social feeding behaviour in *C. elegans* is regulated by antagonistic signals from sensory neurons. Signalling from the ASH and ADL nociceptive neurons promotes social feeding behaviour, whereas signalling from another, as yet unidentified, neuron that expresses *osm-3* suppresses social behaviour by an *npr-1*-independent pathway.

Discussion

Repulsive or stressful environmental conditions may represent important inducers of aggregation behaviour in *C. elegans*. Signalling from either the ASH or the ADL nociceptive neurons is required for social feeding by *npr-1* mutants, suggesting that a repulsive signal stimulates social feeding and related behaviours. However, ASH and ADL neurons are not always essential for social behaviour: a mutation in *osm-3* restores social behaviour to *ocr-2; npr-1* and

odr-4; npr-1 mutants. We suggest that ASH and ADL transmit information about aversive stimuli in the environment to a circuit that is responsible for aggregation, rapid locomotion, and food bordering behaviour. This circuit may also respond to an antagonistic set of neurons that express *osm-3* and inhibit social feeding (Fig. 5c).

Aggregation can be regulated by two different sensory pathways that respond to environmental stress: the *daf-7*/TGF- β dauer pathway¹¹ and the ASH/ADL nociceptive pathway. Similar to aggregation, dauer larva development represents an adaptive response to adverse conditions that is under complex sensory regulation by pheromones, food, and temperature. As double mutants between *daf-7* and *npr-1* (*null*) mutants show enhanced aggregation compared with either mutant alone, it is likely that these two genes use different pathways to inhibit aggregation behaviour.

Food, food acquisition and population density are important regulators of aggregation in a variety of species^{1,2}. Aggregation in nematodes is also regulated by food: well-fed animals from wild social *C. elegans* strains as well as from *npr-1* mutant strains do not aggregate when food is absent. Our results suggest that a nociceptive ASH/ADL stimulus is produced in standard culture conditions, which indicates that it is made at least in part by *E. coli*. It is surprising to think that food would be a source of aversive stimuli, but both *E. coli* and many soil bacteria that form part of the *C. elegans* diet in the wild can kill *C. elegans* under certain conditions^{30–32}.

In mammals, NPY has sedative effects and decreases sensitivity to nociceptive stimuli³³; these effects are reminiscent of our observation that *npr-1* activity suppresses a behaviour induced by noxious stimuli. It will be interesting to study the role of aversive pathways in regulating social behaviour in mammals and other animals. As social feeding is initiated by aversive stimuli, we propose that *C. elegans* aggregation supplies a defence to the animal. Possible advantages to group feeding would include the secretion of enzymes that inactivate bacterial toxins, or the stimulation of dauer formation and dispersal in the next generation. The data in this paper and in the accompanying paper³⁴ suggest that the regulation of social feeding behaviour in *C. elegans* is complex, involving several layers of positive and negative inputs. Such complexity may have evolved as a result of the tension between cooperation and competition that underlies social behaviour, and may be important to ensure that social behaviour is induced only when it offers a selective advantage. □

Methods

Strains, genetics and germline transformation

We grow nematodes at 20 °C under standard conditions³⁵. Double mutants between *npr-1* X and other loci were made by replacing marked balancer chromosomes with chromosomes carrying the mutation of interest. Mutations used have been described previously³⁶. Strains used or generated in this work are listed in Supplementary Information. Germline transformation was carried out as described³⁷. The *lin-15* clone pJMZ (50 ng μ l⁻¹)³⁸ was used as a co-injection marker. Strains injected were either AX215 *odr-4*(n2144) III; *npr-1*(ad609) *lin-15*(n765ts) X or CX4827 *ocr-2*(ak47) IV; *npr-1*(ad609) *lin-15*(n765ts) X. Test DNA was injected at 50 ng μ l⁻¹. At least four transgenic lines were examined for each tested clone.

Behavioural assays

Aggregation and bordering were quantified as described previously, except that 80 animals were used on bacterial lawns of 2-cm diameter⁸. Speed of locomotion was quantified with the DIAS software program⁸. For more detailed information on behavioural assays, see Supplementary Information. The statistical significance of population and single animal behavioural assays was determined using the two-tailed *t*-test and the Mann-Whitney rank sum test, respectively. Error bars represent the s.e.m.

Molecular biology

Details of plasmid construction are available on request. Promoters for *odr-4* constructs were: *odr-10* (AWA); *str-1* (AWB); *odr-3* (AWC, AWB, AWA, ASH, ADF); pT08G3 (ADF); *sra-6* (ASH); *ASI*; *srg-8* (ASK); *sre-1* (ADL, ASJ); pF47C12 (ADL); *str-3* (ASI). *sre-1::odr-4-gfp* and F47C12::*odr-4-gfp*, but not other plasmids, restored social feeding to *odr-4; npr-1* animals. Expression patterns were verified by examining GFP fluorescence. Most functions of ASH are *odr-4*-independent, even those that are mediated by G-protein pathways (C.I.B., unpublished data). Promoters for *ocr-2* constructs were: pT08G3 (ADF);

odr-10 (AWA); *sra-6* (ASH; ASI); pF47C12 (ADL); *odr-3* (AWC, AWB, AWA, ASH, ADF). Laser ablation was used to confirm that rescue of *ocr-2*; *npr-1* social feeding was due to transgene expression in the expected cells. Laser ablation of ADL eliminated rescue by the pF47C12::*ocr-2* transgene. Laser ablation of ASH eliminated rescue by the *sra-6*::*ocr-2* transgene. To aid in the identification of ADL or ASH and to confirm cell death, *ocr-2* transgenic strains also contained either F47C12::GFP (ADL) or *sra-6*::GFP (ASH) markers. Promoters used were described previously^{19,21–23,27,39}.

Laser ablation

Neurons were identified for ablation using DIC optics by a combination of positional and morphological cues^{40,41}. Neurons were killed during the L1 or L2 stages using a laser microbeam⁴¹, and operated animals were assayed within 48 h of the L4 to adult moult. Operated animals were *mls10* V; *npr-1(ad609)* X, which expressed GFP in muscle cells under the *myo-2* promoter. To test aggregation, they were mixed with *npr-1* control animals that did not express GFP.

Received 24 June; accepted 30 September 2002; doi:10.1038/nature01169.

1. Wilson, E. O. *Sociobiology* (Belknap, Harvard Univ., Cambridge, Massachusetts, 1975).
2. Choe, J. C. & Crespi, B. J. (eds) *The Evolution of Social Behavior in Insects and Arachnids* (Cambridge Univ. Press, Cambridge, 1997).
3. Hassell, M. P. & May, R. M. Stability in insect host-parasitoid models. *J. Anim. Ecol.* **42**, 693–736 (1973).
4. Shorrocks, B. & Sevenster, J. G. Explaining local species diversity. *Proc. R. Soc. Lond. B* **260**, 305–309 (1995).
5. McBride, J. M. & Hollis, J. P. The phenomenon of swarming in nematodes. *Nature* **211**, 545–546 (1966).
6. Croll, N. A. *The Behaviour of Nematodes* 18–27 (Edward Arnold, London, 1970).
7. Hodgkin, J. & Doniach, T. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149–164 (1997).
8. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**, 679–689 (1998).
9. Riddle, D. L. & Albert, P. S. *C. elegans II* (eds Riddle, D. L., Blumenthal, T., Meyer, B. & Priess, J. R.) 739–768 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1997).
10. Ren, P. *et al.* Control of *C. elegans* larval development by neuronal expression of a TGF- β homolog. *Science* **274**, 1389–1391 (1996).
11. Thomas, J. H., Birnby, D. A. & Vowles, J. J. Evidence for parallel processing of sensory information controlling dauer formation in *Caenorhabditis elegans*. *Genetics* **134**, 1105–1117 (1993).
12. Schackwitz, W. S., Inoue, T. & Thomas, J. H. Chemosensory neurons function in parallel to mediate a pheromone response in *C. elegans*. *Neuron* **17**, 719–728 (1996).
13. Driscoll, M. & Kaplan, J. C. *C. elegans II* (eds Riddle, D. L., Blumenthal, T., Meyer, B. & Priess, J. R.) 645–678 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1997).
14. Bargmann, C. I. & Mori, I. *C. elegans II* (eds Riddle, D. L., Blumenthal, T., Meyer, B. & Priess, J. R.) 717–738 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1997).
15. Colbert, H. A., Smith, T. L. & Bargmann, C. I. OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in *Caenorhabditis elegans*. *J. Neurosci.* **17**, 8259–8269 (1997).
16. Tobin, D. *et al.* Combinatorial expression of TRPV channel proteins defines their sensory functions and subcellular localization in *C. elegans* neurons. *Neuron* **35**, 307–318 (2002).
17. Caterina, M. J. & Julius, D. The vanilloid receptor: a molecular gateway to the pain pathway. *Annu. Rev. Neurosci.* **24**, 487–517 (2001).
18. Liedtke, W. *et al.* Vanilloid receptor-related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor. *Cell* **103**, 525–535 (2000).
19. Dwyer, N. D., Troemel, E. R., Sengupta, P. & Bargmann, C. I. Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell* **93**, 455–466 (1998).
20. Sawin, E. R., Ranganathan, R. & Horvitz, H. R. *C. elegans* locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. *Neuron* **26**, 619–631 (2000).
21. Sengupta, P., Chou, J. H. & Bargmann, C. I. *odr-10* encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell* **84**, 899–909 (1996).
22. Roayaie, K., Crump, J. G., Sagasti, A. & Bargmann, C. I. The G alpha protein ODR-3 mediates olfactory and nociceptive function and controls cilium morphogenesis in *C. elegans* olfactory neurons. *Neuron* **20**, 55–67 (1998).

23. Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. A. & Bargmann, C. I. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell* **83**, 207–218 (1995).
24. Sambongi, Y. *et al.* Sensing of cadmium and copper ions by externally exposed ADL, ASE, and ASH neurons elicits avoidance response in *Caenorhabditis elegans*. *Neuroreport* **10**, 753–757 (1999).
25. Bargmann, C. I., Thomas, J. H. & Horvitz, H. R. Chemosensory cell function in the behavior and development of *Caenorhabditis elegans*. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 529–538 (1990).
26. Kaplan, J. M. & Horvitz, H. R. A dual mechanosensory and chemosensory neuron in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **90**, 2227–2231 (1993).
27. Troemel, E. R., Kimmel, B. E. & Bargmann, C. I. Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in *C. elegans*. *Cell* **91**, 161–169 (1997).
28. Perkins, L. A., Hedgecock, E. M., Thomson, J. N. & Culotti, J. G. Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **117**, 456–487 (1986).
29. Tabish, M., Siddiqui, Z. K., Nishikawa, K. & Siddiqui, S. S. Exclusive expression of *C. elegans osm-3* kinesin gene in chemosensory neurons open to the external environment. *J. Mol. Biol.* **247**, 377–389 (1995).
30. Garsin, D. A. *et al.* A simple model host for identifying Gram-positive virulence factors. *Proc. Natl Acad. Sci. USA* **98**, 10892–10897 (2001).
31. Troemel, P. S. Influence of bacteria and temperature on the reproduction of *Caenorhabditis elegans* (Nematoda: Rhabditidae) infesting mushrooms (*Agaricus bisporus*). *Nematologica* **37**, 72–82 (1991).
32. Marroquin, L. D., Elyassnia, D., Griffiths, J. S., Feitelson, J. S. & Aroian, R. V. *Bacillus thuringiensis* (Bt) toxin susceptibility and isolation of resistance mutants in the nematode *Caenorhabditis elegans*. *Genetics* **155**, 1693–1699 (2000).
33. Naveilhan, P. *et al.* Reduced antinociception and plasma extravasation in mice lacking a neuropeptide Y receptor. *Nature* **409**, 513–517 (2001).
34. Coates, J. & de Bono, M. Antagonistic pathways in neurons exposed to the body fluid regulate social feeding in *Caenorhabditis elegans*. *Nature* (this issue).
35. Sulston, J. & Hodgkin, J. *The Nematode Caenorhabditis elegans* (ed. Wood, W. B.) 587–606 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1988).
36. Hodgkin, J. C. *C. elegans II* (eds Riddle, D. L., Blumenthal, T., Meyer, B. & Priess, J. R.) 881–1047 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1997).
37. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
38. Clark, S. G., Lu, X. & Horvitz, H. R. The *Caenorhabditis elegans* locus *lin-15*, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. *Genetics* **137**, 987–997 (1994).
39. Sagasti, A., Hobert, O., Troemel, E. R., Ruvkun, G. & Bargmann, C. I. Alternative olfactory neuron fates are specified by the LIM homeobox gene *lim-4*. *Genes Dev.* **13**, 1794–1806 (1999).
40. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
41. Bargmann, C. I. & Avery, L. Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol.* **48**, 225–250 (1995).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank R. Cassada, who first observed aggregation behaviour in a wild isolate of *C. elegans*; D. Madsen and V. Maricq for making the *ocr-2(ak47)* allele available; Y. Zhang and E. Troemel for providing unpublished promoter constructs; A. Davies and S. McIntire for suggesting use of a physical barrier to limit nematode dispersal; A. Fire for expression plasmids; T. Stiernagle and the *Caenorhabditis* Genetics Center for *C. elegans* strains; and M. Hilliard, T. Reader, and J. Coates for suggestions and discussion during the course of this work. This work was supported by the Wellcome Trust, the Howard Hughes Medical Institute, and the Medical Research Council of Great Britain. C.I.B. is an investigator of the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.d.B. (e-mail: debono@mrc-lmb.cam.ac.uk).

A stellar relic from the early Milky Way

N. Christlieb[†], M. S. Bessell[‡], T. C. Beers[§], B. Gustafsson^{*}, A. Korn^{||}, P. S. Barklem^{*}, T. Karlsson^{*}, M. Mizuno-Wiedner^{*} & S. Rossi[¶]

^{*} Hamburger Sternwarte, Gojenbergsweg 112, D-21029 Hamburg, Germany

[†] Department of Astronomy and Space Physics, Uppsala University, Box 524, SE – 75120 Uppsala, Sweden

[‡] Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, Cotter Road, Weston, Australian Capital Territory 2611, Australia

[§] Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA

^{||} Universitäts-Sternwarte München, Scheinerstrasse 1, D-81679 München, Germany

[¶] Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Departamento de Astronomia, 05508-900 São Paulo, Brazil

The chemical composition of the most metal-deficient stars largely reflects the composition of the gas from which they formed. These old stars provide crucial clues to the star formation history and the synthesis of chemical elements in the early Universe. They are the local relics of epochs otherwise observable only at very high redshifts^{1,2}; if totally metal-free ('population III') stars could be found, this would allow the direct study of the pristine gas from the Big Bang. Earlier searches for such stars found none with an iron abundance less than 1/10,000 that of the Sun^{3,4}, leading to the suggestion^{5,6} that low-mass stars could form from clouds above a critical iron abundance. Here we report the discovery of a low-mass star with an iron abundance as low as 1/200,000 of the solar value. This discovery suggests that population III stars could still exist—that is, that the first generation of stars also contained long-lived low-mass objects. The previous failure to find them may be an observational selection effect.

The star HE0107–5240, at coordinates right ascension, RA (2000.0) = 01 h 09 m 29.1 s and declination $\delta = -52^\circ 24' 34''$, is a giant star of the Galactic halo population with apparent magnitude $B = 15.86$. It was found during medium-resolution spectroscopic follow-up observations of candidate metal-poor stars selected from the Hamburg/ESO objective prism survey (HES)^{7,8}. This survey, which covers the entire southern high-galactic-latitude sky to an apparent magnitude limit of $B \approx 17.5$, extends the total survey volume for metal-poor stars in the Galaxy by almost one order of magnitude compared to the total volume explored by previous spectroscopic surveys.

A medium-resolution ($\delta\lambda \approx 0.2$ nm) spectrum of HE0107–5240 was obtained by M.S.B. with the Siding Spring Observatory 2.3-m telescope on 12 November 2001. The Ca II K ($\lambda = 393.4$ nm) line was barely visible in that spectrum, indicating that the star was likely to be extremely metal-deficient. Shortly thereafter, a high-resolution, high signal-to-noise ratio spectrum was obtained with the 8-m Unit Telescope 2 (UT2) of the Very Large Telescope at the European Southern Observatory (ESO), Paranal, Chile.

We derive an effective temperature $T_{\text{eff}} = 5,100 \pm 150$ K for HE0107–5240 by means of broad-band visual and infrared photometry. The absence of Fe II lines, through the Fe I/Fe II ionization equilibrium, constrains the star to have a logarithmic surface gravity of $\log(g) > 2.0$ dex, while main-sequence gravities are excluded by the relative strengths of Balmer lines, and the strength of the Balmer jump seen in the medium-resolution spectrum. Hence we conclude that the star is located on the red-giant branch. By interpolation in a 12-Gyr pre-helium-flash stellar evolutionary track⁹, we estimate the surface gravity of HE0107–5240 to be $\log(g) = 2.2 \pm 0.3$, and its mass $M \approx 0.8M_\odot$. We derive

$[\text{Fe}/\text{H}] = -5.3 \pm 0.2$ for HE0107–5240, where $[A/X] = \log_{10}(N_A/N_X) - \log_{10}(N_A/N_X)_\odot$, and the subscript ' $\odot$ ' refers to the Sun. In the determination of the iron abundance, as well as the Fe I/Fe II ionization equilibrium, we took into account deviations from local thermodynamical equilibrium (LTE) in the formation of iron lines in the atmosphere of the star, which led to a correction of +0.1 dex for the Fe I abundance. The quoted error in the iron abundance includes uncertainties in the derived model atmosphere parameters T_{eff} (resulting in $\delta[\text{Fe}/\text{H}] = \pm 0.20$ dex) and $\log(g)$ ($\delta[\text{Fe}/\text{H}] = \pm 0.02$ dex), the adopted oscillator strengths ($\delta[\text{Fe}/\text{H}] = \pm 0.1$ dex), and in the line-strength measurement uncertainties ($\delta[\text{Fe}/\text{H}] = \pm 0.07$ dex). An LTE analysis, conducted differentially with respect to CD–38° 245 (with $[\text{Fe}/\text{H}] = -3.98$, previously the most iron-deficient giant star known)¹, shows that HE0107–5240 is 1.4 dex more iron-poor. This is in very good agreement with our non-differential LTE value of $[\text{Fe}/\text{H}] = -5.4$ for HE0107–5240.

Extreme iron deficiencies have also been observed in some post-asymptotic giant branch (post-AGB) stars^{10,11}. However, extensive studies have revealed that their observed elemental abundances were altered substantially from their primordial compositions by selective dust depletion and subsequent radiation-driven gas–dust separation^{10,12}. The most extreme cases are HR4049 and HD52961, which have $[\text{Fe}/\text{H}] = -4.8$. Other elements, such as Ca and Mg, are depleted by a similar amount, while the abundances of elements such as C, N, O and Zn are close to the solar values¹¹. For HE0107–5240, on the other hand, we derive an upper limit for the Zn abundance that is significantly below solar ($[\text{Zn}/\text{H}] < -2.7$). We also note that these post-AGB stars occupy a stellar physical parameter space¹¹ that is not shared by HE0107–5240, that is, $T_{\text{eff}} = 6,000$ – $7,600$ K, $\log(g) = 0.5$ – 1.2 . Furthermore, typically many lines in the spectra of post-AGB stars exhibit prominent emission lines owing to a strong stellar wind, while other lines have a complex absorption structure resulting from the presence of circumstellar gas^{11,13}. Such features are not seen in the spectrum of HE0107–5240. Finally, we note that infrared photometry of our star does not reveal any indications of an excess of infrared flux due to hot dust.

We conclude that HE0107–5240 probably formed from a gas cloud with a metal abundance corresponding to $[\text{Fe}/\text{H}] \approx -5.3$. The abundance pattern of elements heavier than Mg can be well fitted by the predicted elemental yields¹⁴ of a 20–25 $M_\odot$ star that underwent a type II supernova explosion, indicating that the gas cloud from which HE0107–5240 formed could have been enriched by such a supernova. (See Table 1.) Alternatively, HE0107–5240 could have formed from a zero-metallicity gas cloud, with its present metallicity being due to accretion of material during repeated passages through the Galactic disk¹⁵.

Table 1 Elemental abundances for HE0107–5240

Element	[X/Fe]
Li	<5.3
C	4.0
N	2.3
Na	0.8
Mg	0.2
Ca	0.4
Ti	–0.4
Ni	–0.4
Zn	<2.7
Sr	<–0.5
Ba	<0.8
Eu	<2.8

Abundance ratios [X/Fe] of HE0107–5240 as derived from a high-resolution, high-S/N UVES spectrum. In our analysis, we used a custom plane-parallel model atmosphere with the most recent atomic and molecular opacity data. Typical errors in the logarithmic abundances, resulting from uncertainties in the stellar parameters and oscillator strengths, are 0.1–0.2 dex. Possible systematic errors are judged to be of the same order of magnitude. The abundances of C, N and Ca have been derived from spectrum synthesis, using the C_2 band at $\lambda = 516.5$ nm, the CN band at $\lambda = 388.3$ nm (assuming a C abundance as listed above), and the Ca II H + K lines, respectively. We measure a carbon isotopic ratio of $^{12}\text{C}/^{13}\text{C} > 30$ from CH A–X lines.

The large overabundances of C and N, and possibly Na, in HE0107–5240 can be explained by either mass transfer from a previously more massive companion during its AGB phase, or else by self-enrichment. The mass-transfer scenario has been proposed as the likely explanation for the so-called metal-poor CH stars¹⁶. Recent model computations of the structure, evolution and nucleosynthesis of low-mass and intermediate-mass stars of zero or near-zero metallicity have shown^{17–19} that such stars undergo extensive mixing episodes at or shortly after the helium-core flash, resulting in a dredge-up of helium burning and CNO-cycle processed material to their surfaces, especially carbon and nitrogen. The surface abundance ratios of CNO, Na and Mg, predicted by Siess *et al.*¹⁹ for stars in the mass range $1\text{--}2M_{\odot}$, agree reasonably well with the abundances observed in HE0107–5240, within the large uncertainties arising from the input physics adopted in the model calculations.

Our derived upper limits for Ba and Sr indicate that HE0107–5240 is not strongly overabundant in neutron-capture elements. In this respect, our star is similar to the extremely metal-poor giant²⁰ CS 22957–27, a star with $[\text{Fe}/\text{H}] = -3.4$ and $[\text{C}/\text{Fe}] = +2.2$, as well as to the recently discovered class of mildly carbon-enhanced metal-poor stars that are underabundant in neutron-capture elements²¹. This abundance pattern implies that helium burning and some CNO burning may have occurred. However, if this scenario is correct, the observations suggest that

the reactions involved have produced little ^{13}C , or that the ^{13}C produced has been consumed by proton captures. We note that if α captures on ^{13}C were dominant, neutrons would have been produced through the reaction $^{13}\text{C}(\alpha, n)^{16}\text{O}$, giving rise to s-process nucleosynthesis.

It was once believed that the lowest-metallicity objects to have formed in the Galaxy (at least those that survived until the present) were the halo globular clusters, such as M92, with $[\text{Fe}/\text{H}] = -2.5$, a factor of 300 times below the solar value. This belief was based on several assumptions, most importantly that star formation in the early Galaxy would have been strongly inhibited at low masses, owing to the difficulty of forming stars from nearly primordial gas without cooling channels arising from heavy elements such as iron. A number of early studies^{22,23} suggested that cooling from molecular species including hydrogen might allow low-mass star formation before the production of heavy metals significantly polluted the interstellar medium, but no examples of stars approaching such low metallicities as HE0107–5240 were then known. This view received support from early objective-prism surveys³ that failed to detect significant numbers of low-mass stars with $[\text{Fe}/\text{H}] < -2.5$. For the past two decades, more extensive surveys have pushed the low-metallicity limit to $[\text{Fe}/\text{H}] = -4.0$, that is, the iron abundance of $\text{CD}\text{--}38^{\circ}245$, but no lower⁴. On the basis of the numbers of metal-poor stars thus far identified in these surveys, which include some 1,000 stars with $[\text{Fe}/\text{H}] < -2.0$, a simple extrapolation of the distribution of metal abundances suggested that if stars with $[\text{Fe}/\text{H}] < -4.0$ existed in the Galaxy, at least a handful should have been found. The dwarf carbon star G77–61 has been suggested to have²⁴ $[\text{Fe}/\text{H}] = -5.5$, if a logarithmic solar Fe abundance of 7.51 (on a scale where the logarithmic abundance of hydrogen is 12) is adopted, as we did for our analysis of HE0107–5240. However, the analysis of the cool ($T_{\text{eff}} = 4,200\text{ K}$)²⁴ dwarf G77–61 is much less certain than the analysis of HE0107–5240, because the spectrum of the former is dominated by molecular bands of carbon that are much stronger than in HE0107–5240. Clearly, the existence of at least one example of a star with an iron abundance as low as $[\text{Fe}/\text{H}] = -5.3$ provides evidence that the limiting metallicity of halo stars may not yet have been reached.

This has implications for the nature of the first mass function (FMF) of early star formation. Although some studies^{5,6} have suggested that metal-poor gas clouds fragment into low-mass objects only above a certain critical metallicity, $Z_{\text{crit}} \approx 10^{-4}Z_{\odot}$ ($[\text{Fe}/\text{H}] = -4$), our discovery provides evidence in favour of other studies suggesting that the FMF included not only very massive stars but also lower-mass stars²⁵. The FMF may also have been bimodal²⁶ and may have included stars with masses around $1M_{\odot}$ as well as around $100M_{\odot}$. Other authors²⁷ have speculated on the possible presence of ancient (presumably extremely low metallicity) objects that could have formed before re-ionization of the early Universe. In this view, the ‘gap’ between the iron abundance of HE0107–5240 and other extremely metal-poor stars may represent the period of reheating and re-ionization, when star formation was strongly suppressed owing to the destruction of the molecular hydrogen that was needed for cooling. As an alternative, it remains possible that the $[\text{Fe}/\text{H}] = -4.0$ ‘limit’ is an artefact caused by the brighter magnitude limits of surveys that have been carried out up to now. Indeed, it may be no coincidence that the fainter magnitude limit of the HES allowed for the detection of HE0107–5240 at a distance of about 11 kpc from the Sun, whereas most previous surveys have been limited to searches for extremely metal-poor stars within the inner halo of the Galaxy. Further tests of this hypothesis are currently targeting the faintest giants from the HES, which extend to 20 or more kpc from the Sun. If additional stars with iron abundances substantially lower than $[\text{Fe}/\text{H}] = -4.0$ are identified, these will provide important new tools with which, for example, we could directly compare observed elemental abundance patterns with the predicted yields of the first generation of type II supernova

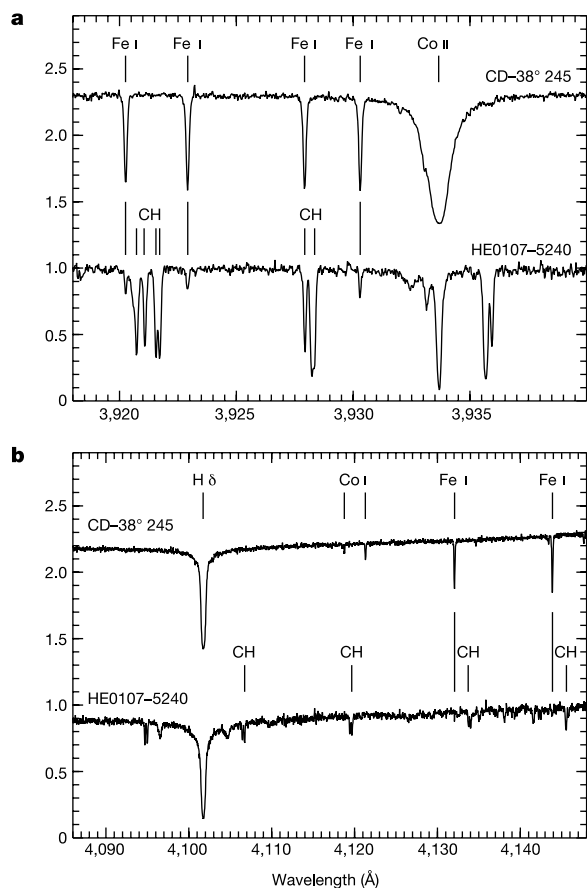


Figure 1 A portion of the spectrum of HE0107–5240, shown compared to the spectrum of $\text{CD}\text{--}38^{\circ}245$, the previously most iron-poor giant star known. Both spectra were obtained with VLT-UT2, and the Ultraviolet-Visual Echelle Spectrograph (UVES). We note the strong molecular CH and C_2 lines and extremely weak lines of Fe I in the spectrum of HE0107–5240. The spectra used in our analysis have a resolution of $R = \lambda/\Delta\lambda = 40,000$, and a signal-to-noise ratio (S/N) of more than 100 per pixel at $\lambda > 400.0\text{ nm}$. The covered wavelength ranges are 329.0–452.0 nm, 478.0–576.0 nm, and 583.0–681.0 nm.

explosions, make a definitive measurement of the primordial lithium abundance (which strongly constrains 'big bang' nucleosynthesis)²⁸, and obtain tighter constraints on the nature of the FME. If any examples of the so-called r-process-enhanced stars are found with $[\text{Fe}/\text{H}] < -4.0$, we may be able to use nucleochronometry²⁹ to obtain a direct estimate of the epoch of first star formation in the Universe, possibly resulting in an improved lower limit for the age of the Universe. □

Received 1 July; accepted 24 September 2002; doi:10.1038/nature01142.

- Norris, J. E., Ryan, S. G. & Beers, T. C. Extremely metal-poor stars. VIII. High-resolution, high signal-to-noise ratio analysis of five stars with $[\text{Fe}/\text{H}] < -3.5$. *Astrophys. J.* **561**, 1034–1059 (2001).
- Cohen, J. G., Christlieb, N., Beers, T. C., Gratton, R. & Carretta, E. Stellar archaeology: A Keck pilot program on extremely metal-poor stars from the Hamburg/ESO survey. I. Stellar parameters. *Astron. J.* **124**, 470–480 (2002).
- Bond, H. E. Where is population III? *Astrophys. J.* **248**, 606–611 (1981).
- Beers, T. C. in *The Third Stromlo Symposium: The Galactic Halo* (eds Gibson, B. K., Axelrod, T. S. & Putman, M. E.) *Astron. Soc. Pacif. Conf. Ser.* **165**, 202–212 (1999).
- Bromm, V., Ferrara, A., Coppi, P. S. & Larson, R. B. The fragmentation of pre-enriched primordial objects. *Mon. Not. R. Astron. Soc.* **328**, 969–976 (2001).
- Schneider, R., Ferrara, A., Natarajan, P. & Omukai, K. First stars, very massive black holes, and metals. *Astrophys. J.* **571**, 30–39 (2002).
- Wisotzki, L. *et al.* The Hamburg/ESO survey for bright QSOs. III. A large flux-limited sample of QSOs. *Astron. Astrophys.* **358**, 77–87 (2000).
- Christlieb, N. *et al.* The stellar content of the Hamburg/ESO survey. I. Automated selection of DA white dwarfs. *Astron. Astrophys.* **366**, 898–912 (2001).
- Yi, S. *et al.* Towards better age estimates for stellar populations: The Y^2 isochrones for solar mixture. *Astrophys. J. Suppl.* **136**, 417–437 (2001).
- Mathis, J. S. & Lamers, H. J. G. L. M. The origin of the extremely metal-poor post-AGB stars. *Astron. Astrophys.* **259**, L39–L42 (1992).
- Van Winckel, H., Waelkens, C. & Waters, L. B. F. M. The extremely iron-deficient "post-AGB" stars and binaries. *Astron. Astrophys.* **293**, L25–L28 (1995).
- Waters, L. B. F. M., Trams, N. R. & Waelkens, C. A scenario for the selective depletion of stellar atmospheres. *Astron. Astrophys.* **262**, L37–L40 (1992).
- Bakker, E. J. *et al.* The optical spectrum of HR 4049. *Astron. Astrophys.* **306**, 924–934 (1996).
- Woosley, S. E. & Weaver, T. A. The evolution and explosion of massive stars. II. Explosive hydrodynamics and nucleosynthesis. *Astrophys. J.* **101**, 181–235 (1995).
- Yoshii, Y. Metal enrichment in the atmospheres of extremely metal-deficient dwarf stars by accretion of interstellar matter. *Astron. Astrophys.* **97**, 280–290 (1981).
- McClure, R. D. & Woodsworth, A. W. The binary nature of the Barium and CH stars. III. Orbital parameters. *Astrophys. J.* **352**, 709–723 (1990).
- Fujimoto, M. Y., Ikeda, Y. & Iben, I. Jr The origin of extremely metal-poor carbon stars and the search for population III. *Astrophys. J.* **529**, L25–L28 (2000).
- Schlattl, H., Salaris, M., Cassisi, S. & Weiss, A. The surface carbon and nitrogen abundances in models of ultra metal-poor stars. *Astron. Astrophys.* (2002) (in the press); also preprint astro-ph/0205326 at (<http://xxx.lanl.gov>).
- Siess, L., Livio, M. & Lattanzio, J. Structure, evolution, and nucleosynthesis of primordial stars. *Astrophys. J.* **570**, 329–343 (2002).
- Norris, J., Ryan, S. G. & Beers, T. C. Extremely metal poor stars. The carbon-rich, neutron capture element-poor object CS 22957–027. *Astrophys. J.* **489**, L169–L172 (1997).
- Aoki, W., Norris, J. E., Ryan, S. G., Beers, T. C. & Ando, H. The chemical composition of carbon-rich, very metal poor stars: a new class of mildly carbon rich objects without excess of neutron-capture elements. *Astrophys. J.* **567**, 1166–1182 (2002).
- Yoshii, Y. & Sabano, Y. Stability of a collapsing pre-Galactic gas cloud. *Publ. Astron. Soc. Jpn* **31**, 505–521 (1979).
- Palla, F., Salpeter, E. E. & Stahler, S. W. Primordial star formation: the role of molecular hydrogen. *Astrophys. J.* **271**, 632–641 (1983).
- Gass, H., Liebert, J. & Wehrse, R. Spectrum analysis of the extremely metal-poor carbon dwarf star G77–61. *Astron. Astrophys.* **189**, 194–198 (1988).
- Yoshii, Y. & Saio, H. Initial mass function for zero-metal stars. *Astrophys. J.* **301**, 587–600 (1986).
- Nakamura, F. & Umemura, M. On the initial mass function of population III stars. *Astrophys. J.* **548**, 19–32 (2001).
- Ostriker, J. P. & Gnedin, N. Y. Reheating of the universe and population III. *Astrophys. J.* **472**, L63–L67 (1996).
- Ryan, S. G., Norris, J. E. & Beers, T. C. The Spite lithium plateau: ultrathin but postprimordial. *Astrophys. J.* **523**, 654–677 (1999).
- Cayrel, R. *et al.* Measurement of stellar age from uranium decay. *Nature* **409**, 691–692 (2001).

Acknowledgements We thank the European Southern Observatory for providing us with reduced UVES spectra. We are grateful to M. Asplund, B. Edvardsson, J. Lattanzio, J. Norris, N. Piskunov, B. Plez, D. Reimers, S.G. Ryan, L. Siess and L. Wisotzki for their contributions and suggestions. N.C. acknowledges a Marie Curie Fellowship granted by the European Commission, and support from Deutsche Forschungsgemeinschaft. T.C.B. acknowledges grants of the US National Science Foundation, S.R. support from FAPESP and CNPq, and the Uppsala group from the Swedish Research Council.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.C. (e-mail: nchristlieb@hs.uni-hamburg.de).

Measurement of the conductance of a hydrogen molecule

R. H. M. Smit*, Y. Noat*†, C. Untiedt*, N. D. Lang‡, M. C. van Hemert§ & J. M. van Ruitenbeek*

*Kamerlingh Onnes Laboratorium, Universiteit Leiden, PO Box 9504, 2300 RA Leiden, The Netherlands

‡IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA

§Leids Instituut voor Chemisch Onderzoek, Gorlaeus Laboratorium, Universiteit Leiden, PO Box 9502, 2300 RA Leiden, The Netherlands

Recent years have shown steady progress towards molecular electronics^{1,2}, in which molecules form basic components such as switches^{3–5}, diodes⁶ and electronic mixers⁷. Often, a scanning tunnelling microscope is used to address an individual molecule, although this arrangement does not provide long-term stability. Therefore, metal–molecule–metal links using break-junction devices^{8–10} have also been explored; however, it is difficult to establish unambiguously that a single molecule forms the contact¹¹. Here we show that a single hydrogen molecule can form a stable bridge between platinum electrodes. In contrast to results for organic molecules, the bridge has a nearly perfect conductance of one quantum unit, carried by a single channel. The hydrogen bridge represents a simple test system in which to understand fundamental transport properties of single-molecule devices.

Here we use a mechanically controllable break junction^{12,13} at low temperatures (4.2 K) to produce pure metallic contacts of atomic size. Figure 1 inset shows a typical conductance curve for a clean Pt contact that was recorded while gradually decreasing the contact size by increasing the piezovoltage (black curve). (We concentrate here on results obtained for Pt wires, but conductance histograms suggest similar behaviour for Pd.) The conductance is expressed in terms of the quantum unit, $G_0 = 2e^2/h$, with e the electron charge and h Planck's constant. The jumps in conductance are the result of sudden atomic rearrangements in response to the applied strain¹⁴. After the conductance has dropped to a value corresponding to a single atom—which for Pt is in the range $1.2–2.3G_0$ —the contact suddenly breaks. In order to extract the common features of these conductance curves, we collect the data of a large series of conductance curves into a conductance histogram. The main panel in Fig. 1 shows such a histogram for Pt contacts (black). It is dominated by a large peak at $1.4–1.8G_0$, which represents the range of conductance values for contacts having a single atom in cross-section (for a recent review, see ref. 15). The histogram drops sharply to zero for lower conductance values, as is typically found for Pt contacts in the absence of adsorbates and impurities.

The character of the conductance curves and the shape of the resulting histogram change markedly when a small quantity of hydrogen gas is admitted to the vacuum vessel (grey curves). The critical amount is difficult to establish, as most hydrogen is expected to condense on the walls of the container (the equilibrium H_2 gas pressure at a temperature of 4.2 K is about 10^{-6} mbar), but the results are not very sensitive to the precise quantity. The peak at the position characteristic for Pt disappears, and a large weight is added in the entire range below that value. On top of this background a distinct peak close to $1G_0$ is found, which grows for larger currents through the contact, while the background is suppressed. For still larger bias voltages, above 200 mV, we recover the histogram for clean Pt. The low-conductance tail and the peak at $1G_0$ reappear

† Present address: Université Denis Diderot (Paris 7), Groupe de Physique des Solides UMR 75 88, 2 place Jussieu, 75251 Paris Cedex 05, France.

upon lowering the bias again. The bias dependence of the histograms is attributed to local heating by the current. For moderate bias the weakly bound physisorbed H_2 is evaporated, and only above 200 mV the chemically bonded hydrogen molecules are removed from the Pt surface.

We concentrate now on the fact that in the presence of hydrogen there appears to be a frequently occurring stable configuration that has a conductance of nearly unity, an example of which is shown in Fig. 1 inset, and which is responsible for the sharp peak just below $1G_0$ in the histogram. We can select this configuration by recording conductance traces of the type shown in the inset, and stop the motion of the electrodes as soon as we find a plateau at $\sim 1G_0$. In order to investigate the structure of the contact on this plateau we use point contact spectroscopy. This technique was originally developed^{16,17} for contacts that are large compared to the atomic scale. The differential conductance, $G_d(V) = dI/dV$, is measured as a function of the d.c. bias voltage. The electrons in the contact are accelerated to an excess energy eV . When this energy reaches that of the main phonon modes of the metal, inelastic scattering results in an enhanced probability for the electrons to scatter back through the contact, which is seen as a drop in $G_d(V)$. This principle has recently been applied to single-atom metallic contacts and chains of atoms^{18,19}, and a theoretical description for localized vibration modes in quantum point contacts has been developed^{20,21}. Vibration modes for individual molecules have been observed before, using inelastic electron tunnelling spectroscopy (IETS) in a low-temperature scanning tunnelling microscope²². Although the principle is similar, the conductance increases at the vibration energy in IETS, whereas it decreases in point contact spectroscopy. Experimentally, a great advantage is the short data acquisition time for a spectrum in our experiment, 10 s compared to 1–10 h for IETS²². This is

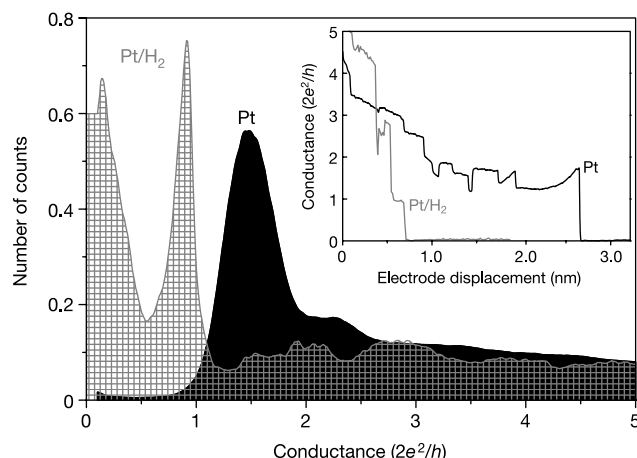


Figure 1 Conductance curves and histograms for clean Pt, and for Pt in a H_2 atmosphere. Inset, a conductance curve for clean Pt (black line) at 4.2 K recorded with a bias voltage of 10 mV, before admitting H_2 gas into the system. About 10,000 similar curves are used to build the conductance histogram shown in the main panel (black), which has been normalized by the area under the curve. After introducing H_2 gas, the conductance curves change qualitatively as illustrated by the grey curve in the inset, recorded at 100 mV. This is most clearly brought out by the conductance histogram (grey; recorded with 140 mV bias). Briefly, the mechanically controllable break-junction technique works as follows. Starting with a macroscopic metal wire, a notch is formed by incision with a knife. The samples are mounted inside a vacuum container and pumped to a pressure below 5×10^{-7} mbar. Next, the system is cooled to 4.2 K in order to attain a cryogenic vacuum. After cooling, the sample wire is broken at the notch by bending of the substrate onto which it has been fixed. The clean, freshly exposed fracture surfaces are then brought back into contact by slightly relaxing the bending. With the use of a piezoelectric element, the displacement of the two electrodes can be finely adjusted to form a stable contact of atomic size. A thick copper finger provides thermal contact to the sample inside the container.

attributed mainly to a lower shot noise level as result of the lower resistance of the junction, and due to the quantum suppression of shot noise for a single-channel contact²³; the lower junction impedance also allows us to work at a higher modulation frequency.

Figure 2 shows the differential conductance and its derivative taken at a $1G_0$ -plateau for the Pt/ H_2 system. We find a pronounced single resonance at about 63.5 mV, symmetrically for both voltage polarities. The energy is much higher than the typical phonon modes for metals, which are found²⁴ between 5 and 25 mV. The width of the resonance is about 14 mV, and is much larger than expected from the thermal and instrumental broadening. A similar large 'intrinsic width' has been observed in IETS²², and it probably results from the short lifetime of the molecular vibration excitations owing to the strong coupling to the metal. We observe a modest variation in the position of the main signal between different experiments, which is probably due to variations in the bonding configuration of the hydrogen to the Pt electrodes. The frequencies obtained from 23 spectra for Pt/ H_2 are shown by the open circles in Fig. 3, having a mean value of 64 mV and a standard deviation of 4 mV.

In order to test the interpretation of the observed resonance, we repeated the experiment using the isotopes D_2 and HD. From 23 spectra for Pt/ D_2 we obtain a distribution of energies shown by open squares in Fig. 3, being centred at 47 mV, while 20 spectra obtained for Pt/HD (filled circles) are found to be centred at 51 mV. Figure 3 inset shows the same distributions with energies scaled by the expected ratios for the vibration energies $\omega_{H_2}/\omega_{D_2} \propto \sqrt{m_{D_2}/m_{H_2}} = \sqrt{2} \approx 1.414$, and $\omega_{H_2}/\omega_{HD} \propto \sqrt{m_{HD}/m_{H_2}} = \sqrt{3/2} \approx 1.225$, confirming our interpretation of the conduction through a hydrogen molecule. Note in particular that this excludes the possibility of conduction through a hydrogen atom, as this would have resulted in a two-peak distribution of frequencies for HD.

Experimental information on the number of conductance modes can be obtained from the fluctuations in $G_d(V)$ as a function of

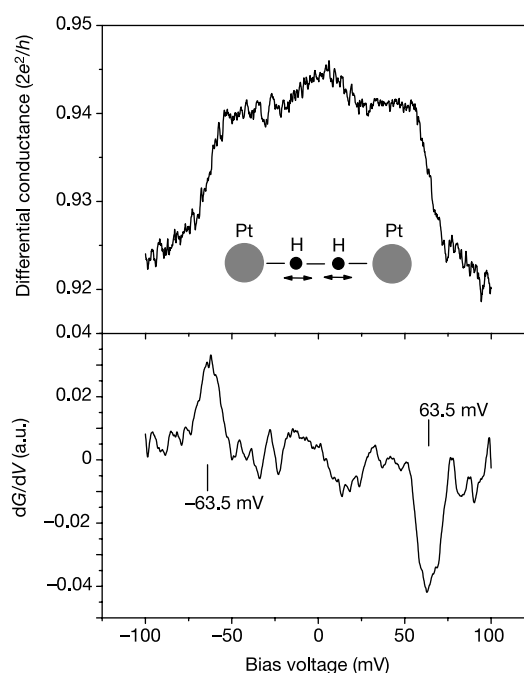


Figure 2 Differential conductance (top) and its derivative (bottom) for a Pt/ H_2 contact taken at a conductance plateau close to $1G_0$. The differential conductance is recorded by a lock-in amplifier using a modulation amplitude between 0.88 and 1.5 mV_{rms} at 7 kHz and a time constant of 10 ms, and the derivative is numerically calculated. A full spectrum is recorded in 10 s.

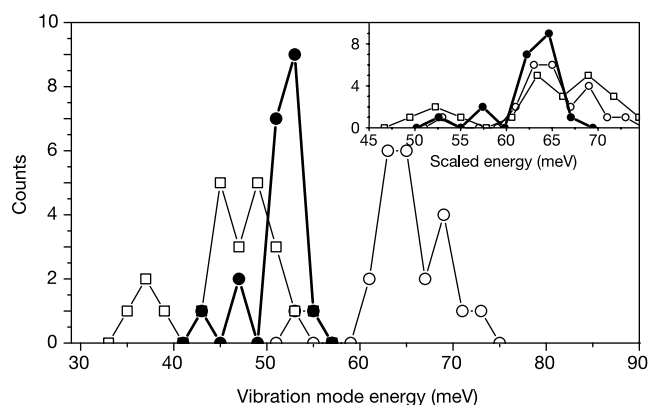


Figure 3 Vibration mode energies obtained from point contact spectra similar to that shown in Fig. 2. Open circles, Pt/H₂; open squares, Pt/D₂; filled circles, Pt/HD. The vertical scale shows the number of spectra with energies within a bin size of 2 meV. The inset shows the same data with the energy axis scaled by the factors expected for the isotope shifts of the hydrogen molecule, $\omega_{\text{H}_2}/\omega_{\text{D}_2} \propto \sqrt{m_{\text{D}_2}/m_{\text{H}_2}} = \sqrt{2} \approx 1.414$ (open squares), and $\omega_{\text{H}_2}/\omega_{\text{HD}} \propto \sqrt{m_{\text{HD}}/m_{\text{H}_2}} = \sqrt{3/2} \approx 1.225$ (filled circles).

voltage by measuring their root-mean-square (r.m.s.) amplitude σ_{GV} following the method described in ref. 25. For a contact with a single conductance channel with transmission probability $T < 1$, the dominant contribution to the fluctuations results from interference of partial waves reflected at the contact itself and those reflected on defects nearby. The wavelength of the electrons changes as a function of the bias voltage, producing random variations in the interference and thus in the conductance. For $T = 1$ the reflection at the contact vanishes, resulting in a suppression of the fluctuations. The peak in the conductance histogram at $0.95G_0$ coincides with a pronounced minimum in σ_{GV} (Fig. 4). From the finite value of σ_{GV} at the minimum²⁵, we extract a value for the transmission $T = 0.97 \pm 0.01$, confirming that the conductance through the molecule is almost entirely carried by one channel. This finding also excludes other configurations for which the conductance would be carried through several parallel channels, and confirms that we have only a single molecule.

Hydrogen is known to bind strongly to a Pt surface, and Pt surfaces catalyse hydrogen dissociation. Little is known about the catalytic activity of this system at 4.2 K, but it is likely that a small energy barrier is present that prevents H₂ dissociation. In our

experiment we cannot rule out the formation of contacts with atomic hydrogen, but the ones that we have been able to fully analyse have a molecular bridge. We have verified the stability of the H₂-bridge configuration using density functional calculations employing the Gaussian 98 program²⁶ with the SDD relativistic effective core basis set²⁷ and the B3LYP functional²⁸. Calculations were performed starting with a linear chain of four Pt atoms, and the results were verified to be insensitive to adding more atoms to the length of the Pt chain. Performing first-principles molecular dynamics calculations, we find that a H₂ molecule bonded to the side of a Pt chain spontaneously moves as a whole into the chain when the bonds between the Pt atoms are being stretched. The energy gain is 2.0 eV, as compared to free H₂ and two Pt₂ chain fragments at infinity. The binding energy of the Pt atoms in the chain is 2.75 eV per bond. H₂ will therefore never move spontaneously into the chain, but first requires an external force to stretch a Pt–Pt bond, which is a way to supply chemical energy to a single bond. In order to obtain the longitudinal vibrational modes of H₂ in a Pt chain, we have explicitly calculated the potential energy curves. For the centre of mass motion of H₂ we obtain an excitation energy of 61.5 meV, using equilibrium bond distances in a linear arrangement of 0.08 nm and 0.21 nm for H–H and Pt–H, respectively. This vibration mode fits very closely the observed excitation energy. The internal vibration mode of H₂ is too high in energy (~ 430 meV) to be observed in the experiment, as the contact becomes unstable at bias voltages above about 200 mV.

The conductance of the hydrogen bridge was calculated using a model of a chain of Pt–H–H–Pt, sandwiched between two ‘jellium’ bulk electrodes. We refer to ref. 29 for more details on the computation method. For the equilibrium bond distances we find a conductance of $0.9G_0$, in close agreement with experiment. We have not attempted to obtain the eigenchannel decomposition of the conductance.

The conduction through the molecule involves mainly the H₂ antibonding states, but the hybridization with the Pt metal states is strong enough to largely fill the gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital. It is surprising that the closed-shell configuration of H₂ permits such strong bonding with Pt, while the molecular character is largely conserved. The latter is shown by the calculated H–H bond distance, which is close to that of the free molecule, and the fact that in the simulations the bridge finally breaks at the Pt–H bond on further stretching. A Pt one-atom contact is expected to have five conductance channels owing to the partially occupied *d* orbitals. It

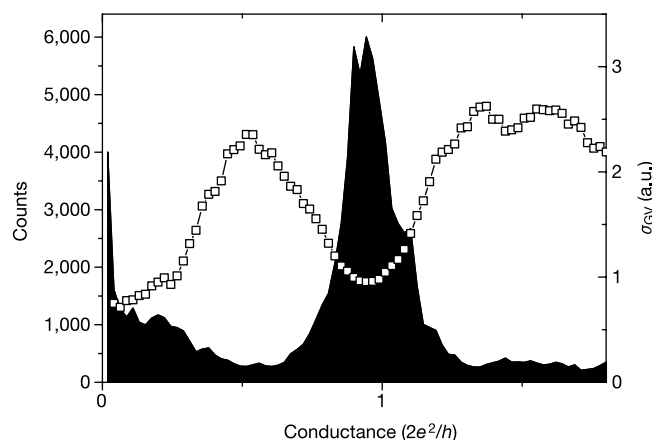


Figure 4 Conductance histogram (black, left axis) and r.m.s. amplitude of the conductance fluctuations σ_{GV} (open squares, right axis) for a Pt/H₂ sample. These data were obtained using 2,000 cycles of contact breaking. The conductance and its derivative were measured with two parallel lock-in amplifiers, detecting the frequencies f and $2f$,

with 140 mV bias voltage and 20 mV modulation amplitude. The derivative signal is used to calculate the average of the conductance fluctuations, σ_{GV} , and each of the points is obtained from the data belonging to one bin of the histogram.

appears that the insertion of a hydrogen molecule has the effect of filtering out one of these five channels, with nearly perfect transmission.

It would be interesting to attempt to extend our technique to more complex molecules with built-in functional groups. Although most organic molecules are expected to have a conductance many orders of magnitude below the quantum unit, our experiments confirm³⁰ that full transmission of a single channel is possible when the coupling to the leads is sufficiently strong. Very recently, two groups have demonstrated conductance through single metal-organic molecules^{31,32}, for which the charge state of the metal ions could even be controlled by a gate electrode. Instead of using mechanical adjustment of the contact size as in our experiments, the size of the metal contacts to the molecule was reduced by exploiting electromigration. The tools needed to study and control electron transport at the single-molecule level are being rapidly developed. □

Received 8 May; accepted 29 August 2002; doi:10.1038/nature01103.

1. Aviram, A. & Ratner, M. (eds) *Molecular Electronics: Science and Technology* (Annals of the New York Academy of Sciences, New York, 1998).
2. Langlais, V. J. *et al.* Spatially resolved tunneling along a molecular wire. *Phys. Rev. Lett.* **83**, 2809–2812 (1999).
3. Gao, H. J. *et al.* Reversible, nanometer-scale conductance transitions in an organic complex. *Phys. Rev. Lett.* **84**, 1780–1783 (2000).
4. Collier, C. P. *et al.* Electronically configurable molecular-based logic gates. *Science* **285**, 391–394 (1999).
5. Reed, M. A., Chen, J., Rawlett, A. M., Price, D. W. & Tour, J. M. Molecular random access memory cell. *Appl. Phys. Lett.* **78**, 3735–3737 (2001).
6. Metzger, R. M. & Cava, M. P. in *Molecular Electronics: Science and Technology* (eds Aviram, A. & Ratner, M.) 95–115 (Annals of the New York Academy of Sciences, New York, 1998).
7. Chen, J., Reed, M. A., Rawlett, A. M. & Tour, J. M. Large on-off ratios and negative differential resistance in a molecular electronic device. *Science* **286**, 1550–1552 (1999).
8. Reed, M. A., Zhou, C., Muller, C. J., Burgin, T. P. & Tour, J. M. Conductance of a molecular junction. *Science* **278**, 252–254 (1997).
9. Kergueris, C. *et al.* Electronic transport through a metal-molecule-metal junction. *Phys. Rev. B* **59**, 12505–12513 (1999).
10. Reichert, J. *et al.* Driving current through single organic molecules. *Phys. Rev. Lett.* **88**, 176804 (2002).
11. Emberly, E. G. & Kirczenow, G. Comment on “First-principles calculation of transport properties of a molecular device”. *Phys. Rev. Lett.* **87**, 269701 (2001).
12. Muller, C. J., van Ruitenbeek, J. M. & de Jongh, L. J. Experimental observation of the transition from weak link to tunneljunction. *Physica C* **191**, 485–504 (1992).
13. van Ruitenbeek, J. M. in *Mesoscopic Electron Transport* (eds Sohn, L. L., Kouwenhoven, L. P. & Schön, G.) 549–579 (Kluwer Academic, Dordrecht, 1997).
14. Rubio, G., Agraït, N. & Vieira, S. Atomic-sized metallic contacts: mechanical properties and electronic transport. *Phys. Rev. Lett.* **76**, 2302–2305 (1996).
15. Agraït, N., Levy Yeyati, A. & van Ruitenbeek, J. M. Quantum properties of atomic-sized conductors. Preprint cond-mat/0208239 at (<http://xxx.lanl.gov>) (2002).
16. Yanson, I. K. Nonlinear effects in the electric conductivity of point junctions and electron-phonon interaction in metals. *Zh. Eksp. Teor. Fiz.* **66**, 1035–1050 (1974); *Sov. Phys. JETP* **39**, 506–513 (1974).
17. Jansen, A. G. M., van Gelder, A. P. & Wyder, P. Point-contact spectroscopy in metals. *J. Phys. C* **13**, 6073–6118 (1980).
18. Untiedt, C., Rubio Bollinger, G., Vieira, S. & Agraït, N. Quantum interference in atomic-sized point-contacts. *Phys. Rev. B* **62**, 9962–9965 (2000).
19. Agraït, N., Untiedt, C., Rubio-Bollinger, G. & Vieira, S. Onset of dissipation in ballistic atomic wires. *Phys. Rev. Lett.* **88**, 216803 (2002).
20. Bonča, J. & Trugman, S. A. Effect of inelastic processes on tunneling. *Phys. Rev. Lett.* **75**, 2566–2569 (1995).
21. Emberly, E. G. & Kirczenow, G. Landauer theory, inelastic scattering and electron transport in molecular wires. *Phys. Rev. B* **61**, 5740–5750 (1999).
22. Stipe, B. C., Rezaei, M. A. & Ho, W. Single-molecule vibrational spectroscopy and microscopy. *Science* **280**, 1732–1735 (1998).
23. van den Brom, H. E. & van Ruitenbeek, J. M. Quantum suppression of shot noise in atomic-sized metallic contacts. *Phys. Rev. Lett.* **82**, 1526–1529 (1999).
24. Khotkevich, A. V. & Yanson, I. K. *Atlas of Point Contact Spectra of Electron-phonon Interactions in Metals* (Kluwer Academic, Dordrecht, 1995).
25. Ludoph, B. & van Ruitenbeek, J. M. Conductance fluctuations as a tool for investigating the quantum modes in atomic-size metallic contacts. *Phys. Rev. B* **61**, 2273–2285 (2000).
26. Frisch, M. J. *et al.* Gaussian 98, Revision A.5 (Gaussian, Inc., Pittsburgh, Pennsylvania, 1998).
27. Andrae, D., Häußermann, U., Dolg, M., Stoll, H. & Preuss, H. Energy-adjusted *ab initio* pseudopotentials for the 2nd and 3rd row transition-elements. *Theor. Chim. Acta* **77**, 123–141 (1990).
28. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
29. Lang, N. D. Resistance of atomic wires. *Phys. Rev. B* **52**, 5335–5342 (1995).
30. Lang, N. D. & Avouris, Ph. Electrical conductance of individual molecules. *Phys. Rev. B* **64**, 125323 (2001).
31. Park, J. *et al.* Coulomb blockade and the Kondo effect in single-atom transistors. *Nature* **417**, 722–725 (2002).
32. Liang, W., Shores, M. P., Bockrath, M., Long, J. R. & Park, H. Kondo resonance in a single-molecule transistor. *Nature* **417**, 725–729 (2002).

Acknowledgements We acknowledge discussions with A. Levy Yeyati and S. K. Nielsen, and we thank D. Bakker and M. Pohlkamp for assistance in the experiments. C.U. and Y.N. were supported by European Community Marie Curie fellowships.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.M.v.R. (e-mail: ruitenbe@phys.leidenuniv.nl).

Relationship between local structure and phase transitions of a disordered solid solution

Ilya Grinberg, Valentino R. Cooper & Andrew M. Rappe

Department of Chemistry and Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, USA

The Pb(Zr,Ti)O₃ (PZT) disordered solid solution is widely used in piezoelectric applications owing to its excellent electro-mechanical properties. Six different structural phases have been observed for PZT at ambient pressure, each with different lattice parameters and average electric polarization. It is of significant interest to understand the microscopic origin of the complicated phase diagram and local structure of PZT^{1–8}. Here, using density functional theory calculations, we show that the distortions of the material away from the parent perovskite structure can be predicted from the local arrangement of the Zr and Ti cations. We use the chemical rules obtained from density functional theory to create a phenomenological model to simulate PZT structures. We demonstrate how changes in the Zr/Ti composition give rise to phase transitions in PZT through changes in the populations of various local Pb atom environments.

PZT is a perovskite ABO₃ alloy with random B-site occupation by either Zr or Ti at all phases and compositions. The six structural phases that have been observed at ambient pressure are an antiferroelectric phase for compositions near PbZrO₃, ferroelectric low-temperature and high-temperature rhombohedral (R) phases for most Zr-rich alloys, a newly discovered monoclinic² (M) ferroelectric phase near 50% Zr/50% Ti ('50/50') composition, a tetragonal (T) ferroelectric phase for Ti-rich alloys, and a cubic paraelectric phase for all compositions at sufficiently high temperature. All phases of the material are generated by distortions from the same high-symmetry cubic parent structure, and are distinguished only by differing lattice parameters and by the directions of the structural distortions which give rise to the average polarization **P** of the material.

We have studied the local structure of PZT using *ab initio* density functional theory (DFT) calculations with the local density approximation (LDA) exchange-correlation functional. To obtain the ground-state structure, we minimize the energy with respect to ionic coordinates, starting with randomized perfect perovskite positions with no symmetry imposed. To represent the bulk nature of the material we use periodic boundary conditions. However, the disorder in the B-cation arrangement makes small, DFT-accessible supercells an inexact representation of the real material. This can be seen from the narrow peaks in the neutron scattering pair-distribution function (PDF) (representing the ensemble average distribution of interatomic distances in the crystal) of the DFT-obtained structure (Fig. 1). Increasing the size of the supercell

quickly becomes computationally intractable, and we therefore take a different approach.

To examine the range of distortion behaviour and to correlate structural motifs with the local environment, we carry out calculations on all possible B-cation arrangements of $4 \times 2 \times 1$ 40-atom supercells at 50/50 composition (Fig. 2) as well as several 60-atom supercells and 30-atom supercells at Ti-rich and Zr-rich compositions. These supercells can be thought of as snapshots of small regions of the real disordered material. As a result, our DFT study identifies distortion patterns and correlations which will also be present in the real disordered material. Predictable dependence of motifs on local structure allows construction of a model that encapsulates the behaviour of the material. Such a model can then be used to study large supercells inaccessible with DFT.

The most energetically important oxygen distortion motif is the contraction (expansion) of Ti–O (Zr–O) bonds by 0.1–0.2 Å in the direction of the neighbouring Zr(Ti) atoms. The BO_6 octahedra have well-defined volumes: 69–71 Å³ for ZrO_6 and 63–65 Å³ for TiO_6 . The oxygen complexes also rotate by 0–5° (ref. 7).

To relate the cation behaviour to local structure, we examine the distortions of the cations from the centres of their oxygen cages. In the perovskite structure, each oxygen participates in four PbO_{12} cages and two B-cation O_6 cages. We define the position of a cage centre as the weighted average of the positions of all the oxygen atoms that participate in the cage, such that the cage centre has an effective charge of the same magnitude as its cation. In an ideal perovskite, the centre of the cage and the metal ion are located at the same point, and polarization is therefore zero. In the actual, lower-symmetry structure such as that shown in Fig. 2, the distortions of the metal ions away from the centre of the oxygen cages generate local polarization and give rise to ferroelectricity.

Our calculations show that Ti and Zr move 0.1–0.3 Å off-centre, with Ti distortions larger and directed more toward one O atom than those of Zr. Pb ions move off-centre by 0.4–0.5 Å, creating four short Pb–O bonds, in agreement with experimental results⁹. Confirming previous interpretation of PZT PDFs⁸, we find that (100), (211) or (110) Pb distortions are preferred over (111), even in the rhombohedral phase where overall **P** is along (111). Unlike the B-cation distortions, the Pb distortion directions are sensitive to the B-cation environment. Pb distortions are toward Ti neighbours and away from Zr neighbours. This is indicated by the difference in the A–B cation distances; the Pb–Ti distances range from 3.27 to 3.4 Å, while Pb–Zr distances range from 3.35 to 3.6 Å. The Pb distortions also conform to the overall polarization as much as possible. The agreement with overall polarization and the Pb atom preference to avoid Zr can either cooperate or compete, as can be

seen from Fig. 2. In the case of Pb atoms 1 and 2 in Fig. 2, both driving forces cooperate. However, in the case of Pb atoms 3 and 4, the local preference to move toward Ti conflicts with the desire to align with overall polarization. This competition results in a compromise, with distortions predominantly along (100) for these Pb atoms. We also find that the alignment of Pb distortions with the average polarization depends on the composition. The scatter of the Pb distortions from the overall polarization vector is 25° for the Zr-rich R phase, 15° for the 50/50 M phase and 10° for the Ti-rich T phase.

Pb distortion directions are sensitive to local structure, whereas the B-cation distortion directions are not. This difference can be understood by considering the differences in the local environment of the A and B cations in the ideal perovskite structure. For Pb atoms, the oxygen nearest-neighbour shell has 12-fold symmetry, but symmetry is broken in the B-cation next-nearest-neighbour shell. Zr is a larger ion than Ti, which means that the purely repulsive interaction between Pb and B cations will be stronger for Pb–Zr than for Pb–Ti. On the other hand, all B-cation environments have six-fold symmetry in the oxygen cage and eight-fold symmetry in the Pb next-nearest-neighbour shell. Therefore, B-cation distortions do not display a strong dependence on local structure and align closely with the Pb atom distortions. In agreement with the suggestion of ref. 8, we therefore identify the Pb distortions as the determining factor for the average structure of the material.

We now examine how the interplay of the dipole interactions and local A–B cation repulsion gives rise to disparate structural phases at different Zr/Ti compositions. The Pb distortions are produced by a hierarchy of interactions. In all PZT phases, Pb tends to move toward (100), (010) and (001) faces in order to form stronger bonds with four of its oxygen neighbours, without closely encountering B cations. (A move by Pb along (111) would be directly toward a B cation, which would incur a high energy cost.) Each Pb has a choice of six such faces. The Pb atoms tend to move toward the face that is closest to the overall polarization direction **P** in order to lower the dipole interaction energy. If this face is Ti-rich (0 or 1 Zr), the Pb will move toward it strongly (perhaps with a small tilt away from the Zr or toward **P**). A neutral face (2 Zr and 2 Ti) permits moderate Pb motion toward it, with a significant tilt. Zr-rich (3 or 4 Zr) faces cause the Pb to tend to move toward a different face to avoid large Pb–Zr repulsion. The relative amounts of Zr-rich, neutral and Ti-rich faces are strongly dependent on Zr/Ti composition (Fig. 3).

In the Ti-rich region of the phase diagram, Ti-rich faces dominate and a T phase (with Pb distortions aligning in one direction) minimizes the dipole alignment energy without raising the repulsion energy. The R phase, with Pb distortions in a variety of directions, minimizes the local repulsion energy just as well as the T phase, but results in a larger dipole interaction energy cost. Therefore the T phase is preferred.

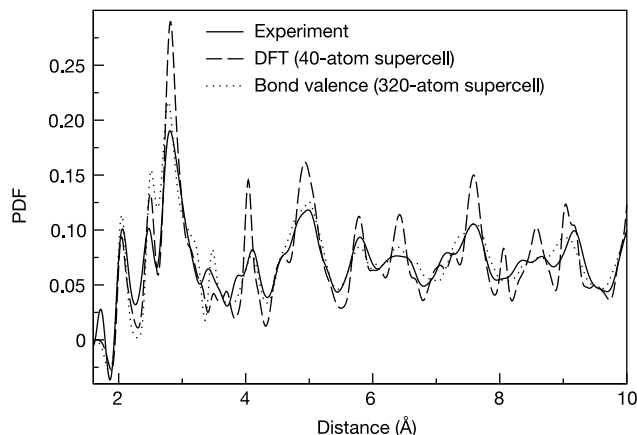


Figure 1 Pair distribution functions (PDFs) for PZT with 50% Zr/50% Ti. Data are from experiment⁸, DFT, and the bond-valence model. Similar agreement with experiment is obtained with the bond-valence model at 40/60 and 60/40 compositions.

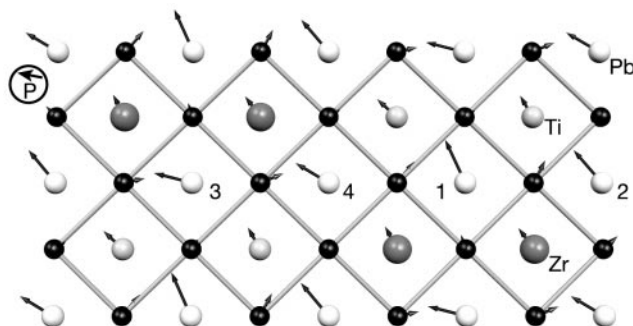


Figure 2 Projection of the $4 \times 2 \times 1$ 50/50 supercell DFT PZT structure on the *x*–*y* plane. The oxygen octahedra are depicted by diamonds, and distortions from the ideal cubic perovskite positions are shown by arrows. Pb atoms are 1/2 unit cell above the plane, and apical O atoms are omitted.

In the Zr-rich region of the phase diagram, Zr-rich faces dominate and T Pb distortions with small angle scatter can minimize the dipole interaction energy only by suffering a high A–B repulsion energy cost. Therefore the R phase is preferred. Here, the Pb atoms distort in a variety of directions toward the available Ti-rich faces to reduce local repulsion cost. However, this large variation in distortion direction raises the dipole interaction energy. Unlike the Ti-rich T phase, the Zr-rich R phase is a disordered compromise structure⁸ where trade-offs are made between the two components of the energy. This explains the difference in angle scatter at various compositions of PZT found by DFT calculations.

At around 50/50 composition, equal amounts of Zr-rich, Ti-rich and neutral faces are present in the material. A concerted distortion in a tetragonal direction can be accommodated by the Pb atoms encountering a Ti-rich or a neutral face. However, the Pb atoms encountering a Zr-rich face must distort in other directions to minimize the A–B repulsion, creating a (211) polarization of the monoclinic phase.

We have created a model to quantify our insights, and to demonstrate that B-cation disorder combined with simple interatomic interactions is sufficient to reproduce and explain the complicated behaviour of PZT as seen by disordered PDFs and compositional phase transitions. The model contains physically intuitive interactions, and can be applied to any perovskite material. We use Ewald summation to compute the long-range Coulomb interaction energy E_{ew} . To simulate bonding we use Brown's rules of valence^{10,11}, which have been tested in many different types of oxides. In the bond valence theory, the bond order of each cation–oxygen bond can be expressed as a power law in the bond distance R_{ij} . Each atom has a desired valence (or total bond order) V_i^0 (+2 for Pb, +4 for Zr and Ti, and –2 for O), so the bond valence energy is

$$E_{\text{bv}} = A_{\text{bv}} \sum_i \left[V_i^0 - \sum_j (R_{ij}/R_{ij}^0)^{-N_{ij}} \right] \alpha_i \quad (1)$$

The j sum runs over all nearest neighbours of the i th atom, R_{ij}^0 and N_{ij} relate bond strength to distance, and A_{bv} and α_i control how stringently the desired bond valences are achieved. As a simple illustration, the ideal perovskite structure would yield valences of 1.5 for Pb, 4.5 for Zr, and 3.6 for Ti. To increase their valences, Pb and Ti must move off-centre and form shorter, stronger bonds. To decrease its valence, Zr must increase its octahedral cage volume.

For better agreement with DFT results, the total energy also includes short-range repulsions and sum rules on short bonds:

$$E = E_{\text{ew}} + E_{\text{bv}} + \sum_{ij} A_{ij} e^{k_{ij}(R_{ij}-D_{ij})} + Q \sum_i (S_i - S_i^0)^\beta \quad (2)$$

A_{ij} , D_{ij} , and k_{ij} control the repulsion strength, to keep the atoms

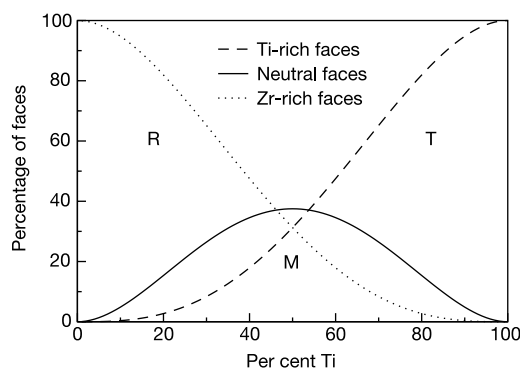


Figure 3 Populations of Ti-rich, Zr-rich and neutral faces in disordered PZT as a function of Ti content. R, M and T label the rhombohedral, monoclinic and tetragonal regions of the phase diagram.

from making unphysically short distances. S_i is the sum of the three shortest Ti–O distances, S_i^0 is the desired value of this sum, and Q and β govern the strength of sum rule.

We parameterize our model to reproduce bond lengths, cage volumes and angle scatter found in DFT calculations, and to give peaks of correct width in the PDF. We then apply the model to study a series of larger $4 \times 4 \times 4$ 320-atom supercells with various random B-cation arrangements and compositions, minimizing the energy with respect to ionic coordinates in the same fashion as in DFT calculations. After minimization we find that the average valence is 4.02 ± 0.05 for Zr and Ti, 2.0 ± 0.08 for O, and 1.98 ± 0.05 for Pb. Our model correctly predicts the compositional phase transitions from rhombohedral at 60/40, to monoclinic at 50/50 to tetragonal at 40/60. This model also yields computed PDFs in excellent agreement with the experimental PDFs (Fig. 1). Thus the interactions we found through DFT calculations, when combined with large disordered cells, result in the rich structure found experimentally in PZT.

As the same physical principles (Ewald summation, Brown's bond-valence interactions, short-range repulsions) govern a wide class of solids, we expect this model to have wide applicability to the study of disordered systems. Expressing the energy of the system in terms of intuitive interatomic interactions allows us to link atomic properties with the statistical distribution of local structural motifs and with macroscopic properties. This multiscale modelling approach may close the gap between the goals of materials design (favourable macroscopic properties) and the techniques used to achieve these goals (typically the variations in the atomic composition of the material). We have shown that a simple statistical interplay of well-known chemical interactions obtained through first-principles calculations, coupled with disorder, can give rise to rich local structure and to complex phase behaviour in a solid solution. □

Methods

In the bond-valence model calculations we use formal charges (+2 for Pb, +4 for Zr and Ti, and –2 for O) for electrostatic interactions. The bond-valence parameters are $A_{\text{bv}} = 80.0$, $R_{\text{Pb-O}} = 2.021$, $R_{\text{Zr-O}} = 1.933$, $R_{\text{Ti-O}} = 1.841$, $N_{\text{Pb-O}} = 5.51$, $N_{\text{Zr-O}} = 6.0$, $N_{\text{Ti-O}} = 5.211$, $\alpha_{\text{Pb}} = 1.25$, $\alpha_{\text{Zr}} = 2.0$, $\alpha_{\text{Ti}} = 2.0$, $\alpha_{\text{O}} = 1.5$, $S_{\text{Ti}}^0 = 5.76$, $\beta = 1.4$, $Q = 20$, $A_{\text{Pb-O}} = 50.0$, $A_{\text{Zr-O}} = 50.0$, $A_{\text{Ti-O}} = 50.0$, $A_{\text{Pb-Zr}} = 175.0$, $A_{\text{Pb-Ti}} = 175.0$, $k_{\text{Pb-O}} = 25.0$, $k_{\text{Zr-O}} = 20.0$, $k_{\text{Ti-O}} = 20.0$, $k_{\text{Pb-Zr}} = 60.0$, $k_{\text{Pb-Ti}} = 60.0$, $D_{\text{Pb-O}} = 2.27$, $D_{\text{Zr-O}} = 1.80$, $D_{\text{Ti-O}} = 1.45$, $D_{\text{Pb-Zr}} = 3.34$, $D_{\text{Pb-Ti}} = 3.23$. A_{bv} and Q are in eV. R_{ij} and D_{ij} are in Å.

Received 24 April; accepted 10 September 2002; doi:10.1038/nature01115.

1. Bellaiche, L., Garcia, A. & Vanderbilt, D. Finite-temperature properties of $\text{Pb}(\text{Zr}_{1-x}\text{Ti}_x)\text{O}_3$ alloys from first principles. *Phys. Rev. Lett.* **84**, 5427–5430 (2000).
2. Noheda, B. et al. A monoclinic ferroelectric phase in the $\text{Pb}(\text{Zr}_{1-x}\text{Ti}_x)\text{O}_3$ solid solution. *Appl. Phys. Lett.* **74**, 2059–2061 (1999).
3. Cohen, R. E. Origin of ferroelectricity in perovskite oxides. *Nature* **358**, 136–138 (1992).
4. Fu, H. & Cohen, R. E. Polarization rotation mechanism for ultrahigh electromechanical response in single-crystal piezoelectrics. *Nature* **403**, 281–283 (2000).
5. Ghosez, P., Cockayne, E., Waghmare, U. V. & Rabe, K. M. Lattice dynamics of BaTiO_3 , PbTiO_3 , and PbZrO_3 : A comparative first-principles study. *Phys. Rev. B* **60**, 836–843 (1999).
6. Saghi-Szabo, G., Cohen, R. E. & Krakauer, H. First-principles study of piezoelectricity in tetragonal PbTiO_3 and $\text{PbZr}_{1/2}\text{Ti}_{1/2}\text{O}_3$. *Phys. Rev. B* **59**, 12771–12776 (1999).
7. Fornari, M. & Singh, D. J. Possible coexistence of rotational and ferroelectric lattice distortions in rhombohedral $\text{PbZr}_x\text{Ti}_{1-x}\text{O}_3$. *Phys. Rev. B* **63**, 092101 (2001).
8. Dmowski, W., Egami, T., Farber, L. & Davies, P. K. in *Fundamental Physics of Ferroelectrics—Eleventh Williamsburg Ferroelectrics Workshop* (ed. Cohen, R. E.) 33–44 (AIP, Woodbury, New York, 2001).
9. Egami, T., Dmowski, W., Akbas, M. & Davies, P. K. in *First-Principles Calculations for Ferroelectrics—Fifth Williamsburg Workshop* (ed. Cohen, R. E.) 1–10 (AIP, Woodbury, New York, 1998).
10. Brese, N. E. & O'Keeffe, M. Bond-valence parameters for solids. *Acta Crystallogr. B* **47**, 192–197 (1991).
11. Brown, I. D. in *Structure and Bonding in Crystals II* (eds O'Keeffe, M. & Navrotsky, E.) 1–30 (Academic, New York, 1981).

Acknowledgements We thank T. Egami, W. Dmowski and P. K. Davies for discussions and for sharing PDF data. This work was supported by the Office of Naval Research and the National Science Foundation. Computational support was provided by the High-Performance Computing Modernization Office of the Department of Defense and the Center for Piezoelectric Design. A.M.R. thanks the Camille and Henry Dreyfus Foundation for support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.M.R. (e-mail: rappe@sas.upenn.edu).

High tensile ductility in a nanostructured metal

Yinmin Wang*, Mingwei Chen†, Fenghua Zhou† & En Ma*

* Department of Materials Science and Engineering,

† Department of Mechanical Engineering, The Johns Hopkins University, Baltimore, Maryland 21218, USA

Nanocrystalline metals—with grain sizes of less than 100 nm—have strengths exceeding those of coarse-grained and even alloyed metals^{1,2}, and are thus expected to have many applications. For example, pure nanocrystalline Cu (refs 1–7) has a yield strength in excess of 400 MPa, which is six times higher than that of coarse-grained Cu. But nanocrystalline materials often exhibit low tensile ductility at room temperature, which limits their practical utility. The elongation to failure is typically less than a few per cent; the regime of uniform deformation is even smaller^{1–7}. Here we describe a thermomechanical treatment of Cu that results in a bimodal grain size distribution, with micro-metre-sized grains embedded inside a matrix of nanocrystalline and ultrafine (<300 nm) grains. The matrix grains impart high strength, as expected from an extrapolation of the Hall–Petch relationship. Meanwhile, the inhomogeneous microstructure induces strain hardening mechanisms^{8–11} that stabilize the tensile deformation, leading to a high tensile ductility—65% elongation to failure, and 30% uniform elongation. We expect that these results will have implications in the development of tough nanostructured metals for forming operations and high-performance structural applications including microelectromechanical and biomedical systems.

The pure copper used in this model study is very ductile in its annealed and coarse-grained form. It has an elongation to failure as large as 70% (curve A, Fig. 1), but a low yield strength (σ_y). Strengthening through heavy cold work results in a material with a tensile curve that peaks immediately after yielding (curve B, Fig. 1). Such a trend of strengthening accompanied by a loss of ductility is generally true for Cu and other metals processed in various ways^{1–7,12}, as summarized in Fig. 2.

Our processing starts by rolling the Cu at liquid nitrogen temperature to a high value of percentage cold work (CW; see Methods). The resulting material has a typical heavily deformed microstructure with high densities of dislocations in nanoscale networks^{13,14}, with some resolvable grains all below 200 nm in size¹⁴. The use of the low temperature suppresses dynamic recovery, allowing the density of the accumulated dislocations to reach a higher steady-state level than that achievable at room temperature. Transmission electron micrographs of the microstructures after recovery annealing and recrystallization are shown in Fig. 3a, b. The recrystallized grains have well-defined, high-angle boundaries (Fig. 3b). The majority of the grains are in the nanocrystalline to ultrafine range. Meanwhile, abnormal grain growth (secondary recrystallization¹⁵) has started to produce a volume fraction (~25%) of coarser (1–3 μ m) grains, some of which contain annealing twins.

The engineering stress–strain curves corresponding to the microstructures in Fig. 3 are compared in Fig. 1. After 93% CW at liquid-nitrogen temperature (curve C), the σ_y is much higher than that of the room-temperature-rolled Cu (95% CW; curve B). The elongation to failure is also significantly larger. After the 180 °C annealing, the σ_y decreased slightly for the recovered and partially recrystallized grains, and the elongation to failure increased further (curve D). Such concurrent strengthening and toughening (in terms of post-necking elongation), as observed for both curves C and D when compared with curve B, can be attributed to the

nanocrystalline/ultrafine grains that reduce the size of nucleating flaws and increase the resistance to crack propagation, leading to a higher fracture stress¹. Micrographs of the fracture surfaces (not shown) indicate that ductile fracture through the nucleation and coalescence of extremely fine microvoids was promoted¹⁴. The radial unstable crack growth was delayed, and the stabilizing triaxial stress state was maintained to larger strains¹⁶.

A marked improvement in uniform elongation was found concurrent with pronounced strain hardening, without sacrificing much of the strength, in material with the bimodal microstructure shown in Fig. 3b. The resultant stress–strain curve is shown in Fig. 1 (curve E). The large densities of defects and cold-work energy stored during processing at liquid nitrogen temperature allowed copious nucleation during recrystallization at a low temperature, so that the vast majority of the matrix grains are kept in the nanocrystalline/ultrafine grain regime to help maintain the high strength of the ‘composite’ material. Meanwhile, the grains with sufficiently large sizes obtained through secondary recrystallization, at a volume fraction of 25%, produced pronounced strain hardening to sustain the useful uniform deformation to large strains. Note that one should not restore strain hardening by allowing uniform growth of all grains or a large fraction of excessively grown grains, both of which would cause an additional unwanted drop in σ_y .

The strain hardening is needed because the onset of localized deformation (necking instability, peak in Fig. 1) in tension is governed by the Considère criterion^{17–19}

$$\left(\frac{\partial \sigma}{\partial \varepsilon}\right)_{\varepsilon} \leq \sigma \quad (1)$$

where σ and ε are true stress and true strain, respectively. The nanocrystalline and ultrafine matrix grains tend to lose the work hardening (left-hand side of equation (1)) quickly on deformation owing to their very low dislocation storage efficiency inside the tiny grains, especially when in presence of dynamic recovery²⁰. Such a high-strength material is therefore prone to plastic instability (early necking), severely limiting the desirable uniform elongation unless larger grains of appropriate sizes and volume fractions are present.

Our strategy is to efficiently use the moderate population of the larger grains to achieve a strain hardening rate much higher than that which would be expected from curve A (Fig. 1, for uniform deformation under uniaxial tension). We do this by taking advantage of the following three factors. First, these confined grains deformed in the inhomogeneous microstructure are under multi-axial stress states. There are complex strain paths and triaxial

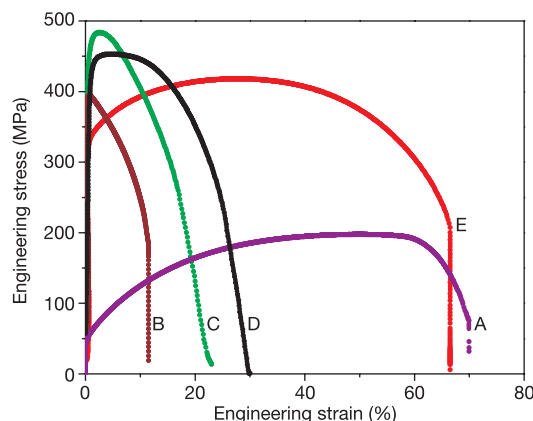


Figure 1 Engineering stress–strain curves for pure Cu. Curve A, annealed, coarse-grained Cu; B, room temperature rolling to 95% cold work (CW); C, liquid-nitrogen-temperature rolling to 93% CW; D, 93% CW + 180 °C, 3 min.; and E, 93% CW + 200 °C, 3 min. Note the coexisting high strength and large uniform plastic strain as well as large overall percentage elongation to failure for curve E.

strain components, with very large strain gradients (see Supplementary Information). It is known that a complex stress state, complicated straining patterns and dislocation configurations, and high densities of geometrically necessary dislocations are all beneficial for promoting grain refinement (or dislocation storage and strain hardening). For example, equal channel angular pressing (ECAP) uses similar conditions to make nanostructured metals in an efficient way²¹. For a metal such as Cu, the non-uniform deformation over a length scale of the order of a few micrometres (Fig. 3b) is the realm of strain-gradient plasticity theory¹¹, which predicts significant strain hardening owing to an excessively large number of geometrically necessary dislocations that are forced to be present to accommodate the large strain gradient.

Second, $\{112\} \{111\}$ twinning, as shown in Fig. 4a and the selected-area electron diffraction pattern in the upper left inset, was observed unexpectedly after straining for 6% inside most of these larger grains. Deformation twins have not been observed for Cu before except at high strain rates or low, sub-ambient temperatures, and activation of such twins is known to require high stresses, especially when the grain sizes are small^{8,9,22,23}. Additional observations by transmission electron microscopy (for example, the high-resolution image in Fig. 4a lower right inset) show twin boundaries located preferentially near the protrusions of the surrounding nanocrystalline/ultrafine grains into the softer large grains, suggesting twinning initiation presumably due to stress concentration. The activation of the twinning mechanism suggests that these constrained larger grains plastically deform at high stresses, consistent with the high strength observed. In terms of enhancing strain hardening, twinning is known to be highly effective in conventional Cu (refs 8, 9), owing to dislocation pileups at the twin boundaries (Fig. 4a). In nanocrystalline Cu, the interfaces generated between twinned segments can act as strong barriers to dislocation motion¹⁰.

Third, the larger (softer) grains accommodate strains preferentially (see Supplementary Information). When the overall uniform elongation reached $\sim 30\%$ (peak of curve E in Fig. 1), these larger grains had accumulated large numbers of twin boundaries, dislocations and subgrain boundaries such that the microstructure was refined to a level similar to the nanocrystalline/ultrafine grained matrix, Fig. 4b). Afterwards, the post-necking elongation is similar

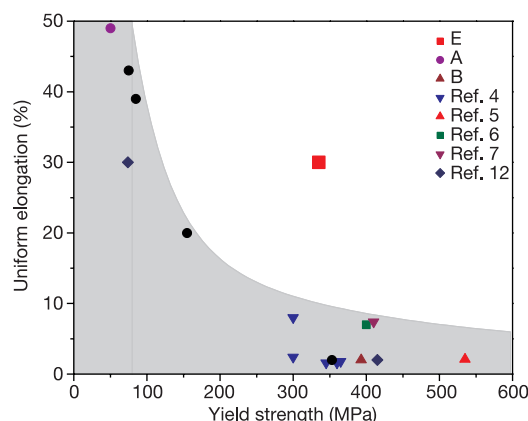


Figure 2 Representative tensile properties of pure Cu. The data are for Cu of conventional, ultrafine and nanocrystalline grain sizes^{4–7,12}, and after cold rolling to various degrees of CW from our own tests (filled black circles). Data points E, A and B are from the corresponding curves in Fig. 1. Uniform elongation up to the peak in the engineering stress–strain curve is compared here, not only because it is a desirable property but also because the overall percentage elongation to failure is often dominated by localized deformation (necking) whose magnitude depends on the gauge length used in the different experiments. The ‘tough Cu’ developed here (E) is clearly separated from the general trend.

to that discussed for the sample annealed at 180 °C (curve D in Fig. 1). Overall, the nanostructured Cu of curve E (Fig. 1), when compared with the coarse grained starting material, represents an elevation of σ_y by a factor of 5–6 while maintaining comparable elongation to failure. The simultaneous high strength and ductility, especially the very large uniform deformation at the elevated strength, results in a notable gain in toughness (the area under the stress–strain curve)¹⁶. This is what sets this material apart from all previous treated copper materials, as demonstrated in Fig. 2.

To establish the reproducibility, three additional samples with or without the ECAP step were processed through similar CW and heat treatment. In all cases, coexisting high strength and ductility were observed. Further annealing beyond that shown in Fig. 3b caused additional grain growth and larger uniform elongation, but with a large decrease in σ_y and no gain in overall ductility owing to the decrease of the post-peak elongation (compare curve A with curve D, Fig. 1). Attempts to start with the room-temperature-rolled Cu only managed a σ_y of ~ 100 MPa when elongation to failure reached $\sim 50\%$. This emphasizes the importance of the step at liquid nitrogen temperature, which stores large cold work energy that leads to a lower recrystallization temperature (compared with room-temperature rolling, the calorimetric recrystallization/grain growth peak temperature decreased by 60 °C) and favours copious nucleation over growth. This makes it possible to achieve the

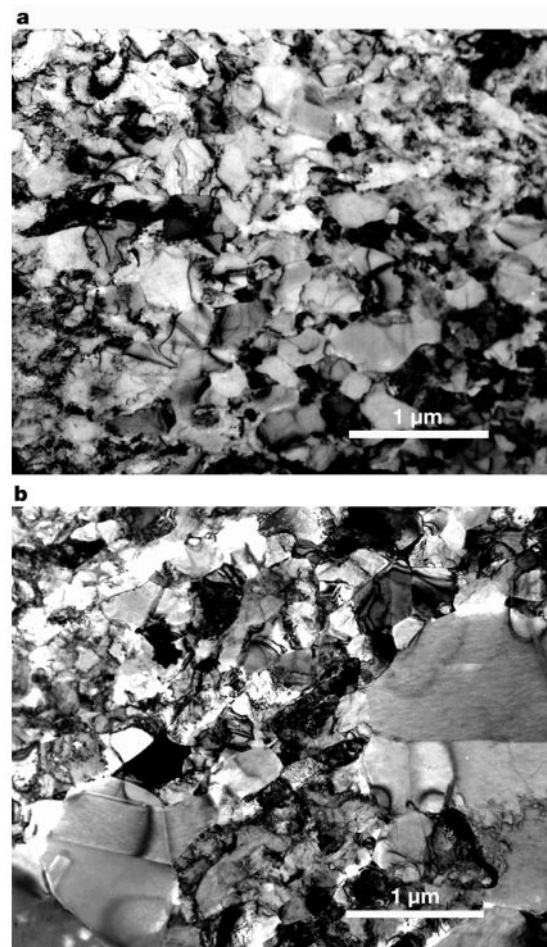


Figure 3 Transmission electron micrographs showing the evolution of the Cu microstructure. Panels **a** and **b** show the samples used to obtain the curves D and E in Fig. 1, respectively. After annealing at 180 °C for 3 min (**a**), recovery has occurred, and the dislocation density is much reduced. The vast majority of the grains are in the nanocrystalline/ultrafine range, with some recrystallized regions. Heat-treating at 200 °C for 3 min led to full recrystallization followed by secondary recrystallization (**b**).

nanocrystalline/ultrafine grained matrix structure through recrystallization, thus affording room for tailoring the microstructure through controlled secondary recrystallization¹⁵.

Our approach does not use uniform nanocrystalline grains, which have to rely on grain boundary deformation mechanisms (such as grain boundary sliding) to contribute significantly to ductility and stabilize the plastic deformation through large increases in strain rate sensitivity^{6,24–26}. Experimental data so far (for example, Fig. 2) indicate that at ambient temperature the increase in ductility provided by grain boundary sliding in small grains is either insufficient to compensate for the loss of dislocation controlled ductility, or concurrent with a much reduced strength^{2,6}.

Our idea of improving strain hardening may be used to derive good ductility from other nanocrystalline materials, where abnormal grain growth is often observed. For example, after heating to a moderate temperature nanocrystalline nickel was reported to exhibit a bimodal microstructure and pronounced strain hardening under certain deformation conditions^{25–27}. In addition to achieving a combination of strength and ductility, our thermomechanical approach to the processing of bulk samples is also simpler than those processes required to produce uniform nanocrystalline grains; the latter processes are not only difficult or expensive to implement, but also are difficult to keep free of artefacts such as porosity and

impurities. In fact, fracture due to sample flaws after consolidation or deposition^{2–4}, together with plastic flow instabilities such as necking and shear banding²⁸, are responsible for the very limited strain to failure observed so far in nanocrystalline materials. □

Methods

Pure Cu (99.99%) bar from a commercial source (ESPI) was first processed by severe cold rolling, with liquid-nitrogen-temperature (LNT) cooling of the workpiece between consecutive rolling passes (–150 °C and –100 °C before and after each pass, respectively). The degree of LNT deformation is defined using per cent cold work²⁹, %CW = $(S_0 - S)/S_0 \times 100\%$, where S_0 and S are the cross-sectional areas before and after rolling. Some samples were processed through eight passes of equal channel angular pressing (ECAP) at room temperature before rolling, while others were subjected directly to LNT rolling. The ECAP step made no obvious difference to the eventual microstructure and properties, as the very large cold work energy stored at LNT controls the subsequent recrystallization behaviour. Microhardness, as well as the recrystallization temperature and enthalpy storage measured in a calorimetric scan, levels off after about 90% CW.

For mechanical property measurements, all the samples were cut and polished to a cross-section of 1 mm × 1.8 mm, and a gauge length of 5 mm (previously reported tensile tests of nanostructured metals typically used a gauge length in the range of 1–5 mm; refs 2, 5, 6, 25, 26). Uniaxial tensile tests were conducted at room temperature at an initial quasi-static strain rate of $5 \times 10^{-4} \text{ s}^{-1}$.

Received 23 April; accepted 12 September 2002; doi:10.1038/nature01133.

- Koch, C. C., Morris, D. G., Lu, K. & Inoue, A. Ductility of nanostructured materials. *Mater. Res. Soc. Bull.* **24**, 54–58 (1999).
- Weertman, J. R. et al. Structure and mechanical behavior of bulk nanocrystalline materials. *Mater. Res. Soc. Bull.* **24**, 44–50 (1999).
- Sanders, P. G., Youngdahl, C. J. & Weertman, J. R. The strength of nanocrystalline metals with and without flaws. *Mater. Sci. Eng. A* **234–236**, 77–82 (1997).
- Sanders, P. G., Eastman, J. A. & Weertman, J. R. Elastic and tensile behavior of nanocrystalline copper and palladium. *Acta Mater.* **45**, 4019–4025 (1997).
- Lergos, M., Elliott, B. R., Rittner, M. N., Weertman, J. R. & Hemker, K. J. Microsample tensile testing of nanocrystalline metals. *Phil. Mag.* **80**, 1017–1026 (2000).
- Valiev, R. Z., Alexandrov, I. V., Zhu, Y. T. & Lowe, T. C. Paradox of strength and ductility in metals processed by severe plastic deformation. *J. Mater. Res.* **17**, 5–8 (2002).
- Gertsman, V. Y., Valiev, R. Z., Akhmediev, N. A. & Mishin, O. V. Deformation behavior of ultrafine-grained materials. *Mater. Sci. Forum* **225–227**, 739–744 (1996).
- Asgari, S., El-Danaf, E., Kalidindi, S. R. & Doherty, R. D. Strain hardening regimes and microstructural evolution during large strain compression of low stacking fault energy fcc alloys that form deformation twins. *Metall. Mater. Trans. A* **28**, 1781–1795 (1997).
- Andrade, U., Meyers, M. A., Vecchio, K. S. & Chokshi, A. H. Dynamic recrystallization in high-strain, high-strain-rate plastic deformation of copper. *Acta Metall. Mater.* **42**, 3183–3195 (1994).
- Youngdahl, C. J., Weertman, J. R., Hugo, R. C. & Kung, H. H. Deformation behavior in nanocrystalline copper. *Scripta Mater.* **44**, 1475–1478 (2001).
- Gao, H., Huang, Y., Nix, W. D. & Hutchinson, J. W. Mechanism-based strain gradient plasticity-I. Theory. *J. Mech. Phys. Solids* **47**, 1239–1263 (1999).
- Lu, L., Wang, L. B., Ding, B. Z. & Lu, K. High-tensile ductility in nanocrystalline copper. *J. Mater. Res.* **15**, 270–273 (2000).
- Hughes, D. A. & Hansen, N. High angle boundaries formed by grain subdivision mechanisms. *Acta Mater.* **45**, 3871–3886 (1997).
- Wang, Y. M., Ma, E. & Chen, M. W. Enhanced tensile ductility and toughness in nanostructured Cu. *Appl. Phys. Lett.* **80**, 2395–2397 (2002).
- Humphreys, F. J. & Hatherly, M. *Recrystallization and Related Annealing Phenomena*, 1st edn 314 (Pergamon, New York, 1995).
- Hertzberg, R. W. *Deformation and Fracture Mechanics of Engineering Materials*, 3rd edn 89, 392 (Wiley and Sons, New York, 1989).
- Hart, E. W. Theory of the tensile test. *Acta Metall.* **15**, 351–355 (1967).
- Dieter, G. E. *Mechanical Metallurgy*, 3rd edn 290 (McGraw-Hill, Boston, 1986).
- Jia, D. et al. Deformation behavior and plastic instabilities in ultrafine-grained Ti. *Appl. Phys. Lett.* **79**, 611–613 (2001).
- Wang, Y. M. & Ma, E. Strain hardening, strain rate sensitivity, and ductility of nanoconstructed metals. *Mater. Sci. Eng. A* (in the press).
- Valiev, R. Z., Islamgaliev, R. K. & Alexandrov, I. V. Bulk nanostructured materials from severe plastic deformation. *Prog. Mater. Sci.* **45**, 103–189 (2000).
- Huang, J. Y., Wu, Y. K. & Ye, H. Q. Deformation structures in ball milled copper. *Acta Mater.* **44**, 1211–1221 (1996).
- Blewitt, T. H., Coltman, R. R. & Redman, J. K. Low-temperature deformation of copper single crystals. *J. Appl. Phys.* **28**, 651–660 (1957).
- Lu, L., Sui, M. L. & Lu, K. Superplastic extensibility of nanocrystalline copper at room temperature. *Science* **287**, 1463–1465 (2000).
- McFadden, S. X., Mishra, R. S., Valiev, R. Z., Zhilyaev, A. P. & Mukerjee, A. K. Low-temperature superplasticity in nanostructured nickel and metal alloys. *Nature* **398**, 684–686 (1999).
- McFadden, S. X., Zhilyaev, A. P., Mishra, R. S. & Mukerjee, A. K. Observation of low-temperature superplasticity in electroplated ultrafine grained nickel. *Mater. Lett.* **45**, 345–349 (2000).
- Hibbard, G. D., McCrea, J. L., Palumbo, G., Aust, K. T. & Erb, U. An initial analysis of mechanisms leading to late stage abnormal grain growth in nanocrystalline Ni. *Scripta Mater.* **47**, 83–87 (2002).
- Wei, Q. M., Jia, D., Ramesh, K. T. & Ma, E. Evolution and microstructure of shear bands in nanostructured Fe. *Appl. Phys. Lett.* **81**, 1240–1242 (2002).
- Callister, W. D. Jr. *Materials Science and Engineering*, 3rd edn 167 (Wiley and Sons, New York, 1994).

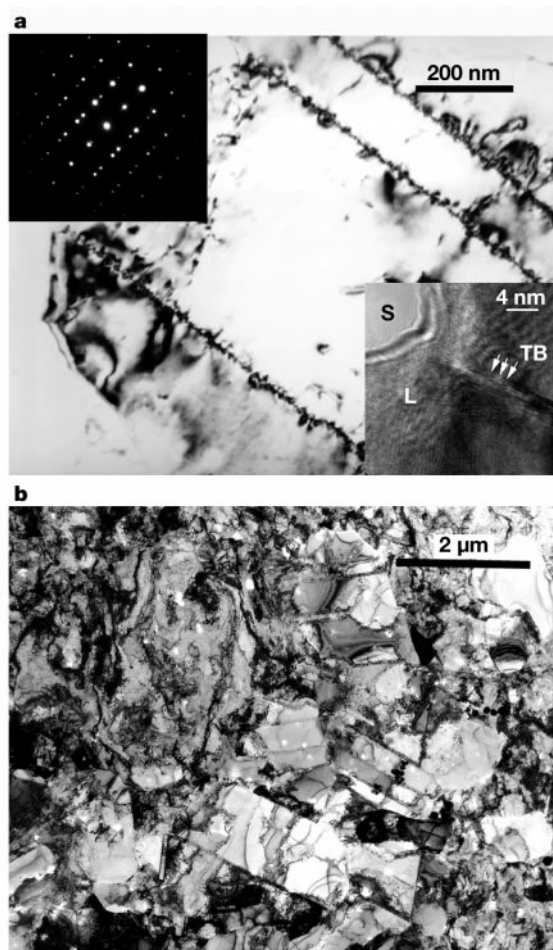


Figure 4 Transmission electron micrographs of Cu after different tensile strains. The Cu sample is that shown in Fig. 3b. **a**, after 6% plastic strain; the upper left inset shows the selected-area electron diffraction pattern, and the lower right inset shows the high-resolution image of the boundary region between a larger, micrometre-sized grain (L) and one of the surrounding much smaller ultrafine grains (S). A twin boundary (TB) is seen near the 'tip' of the S protrusion into L, where twinning is initiated. **b**, After the maximum uniform strain of ~30%.

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements We thank G. Xu, H. Gao, S. X. Mao and D. van Heerden for discussions. This work was supported by the US National Science Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.M. (e-mail: ema@jhu.edu).

Variable effects of nitrogen additions on the stability and turnover of soil carbon

Jason C. Neff^{*,†}, Alan R. Townsend^{†,‡,§}, Gerd Gleixner^{||}, Scott J. Lehman[§], Jocelyn Turnbull[§] & William D. Bowman[‡]

^{*} Earth Surface Processes Team, Geologic Division, US Geological Survey, MS 980, Denver Federal Center, Denver, Colorado 80225, USA

[‡] Environmental, Population and Organismic Biology; and [§] Institute for Arctic and Alpine Research, University of Colorado at Boulder, Boulder, Colorado 80309, USA

^{||} Max Planck Institute for Biogeochemistry, Jena, Germany

[†] These authors contributed equally to this work

Soils contain the largest near-surface reservoir of terrestrial carbon¹ and so knowledge of the factors controlling soil carbon storage and turnover is essential for understanding the changing global carbon cycle. The influence of climate on decomposition of soil carbon has been well documented^{2,3}, but there remains considerable uncertainty in the potential response of soil carbon dynamics to the rapid global increase in reactive nitrogen (coming largely from agricultural fertilizers and fossil fuel combustion). Here, using ¹⁴C, ¹³C and compound-specific analyses of soil carbon from long-term nitrogen fertilization plots, we show that nitrogen additions significantly accelerate decomposition of light soil carbon fractions (with decadal turnover times) while further stabilizing soil carbon compounds in heavier, mineral-associated fractions (with multidecadal to century lifetimes). Despite these changes in the dynamics of different soil pools, we observed no significant changes in bulk soil carbon, highlighting a limitation inherent to the still widely used single-pool approach to investigating soil carbon responses to changing environmental conditions. It remains to be seen if the effects observed here—caused by relatively high, short-term fertilizer additions—are similar to those arising from lower, long-term additions of nitrogen to natural ecosystems from atmospheric deposition, but our results suggest nonetheless that current models of terrestrial carbon cycling do not contain the mechanisms needed to capture the complex relationship between nitrogen availability and soil carbon storage.

Human activity now fixes more atmospheric N₂ into biologically available forms each year than all natural processes combined⁴,

causing a wide range of cascading environmental responses, including a possible sink for excess atmospheric CO₂ through stimulation of plant growth in N-limited ecosystems⁵. However, on average, soils contain three times as much C as does terrestrial vegetation¹. Thus if changes in N availability alter soil C turnover, net C sinks from increased plant growth could be significantly enhanced or reduced, depending on the direction of the soil responses. Unfortunately, considerable uncertainty remains concerning the relationship between N availability and decomposition processes. Additions of N and/or natural variation in N concentrations have led to increases, decreases or no change in observed decomposition rates. This is true for field studies of litter and soil organic matter (SOM) decomposition, as well as for laboratory experiments^{6–10}.

In part, these varied results occur because decomposition studies are difficult to carry out and interpret. Soils are highly complex media with a diversity of substrates that vary in both the energy required for their breakdown and in their total N content. As a result, additions of N to soils could increase decomposition rates of some SOM fractions while simultaneously decreasing rates for other fractions. Past results suggesting widely varied—or no—responses of decomposition to N inputs may be due to varied and compensatory responses of SOM fractions that are difficult to detect with traditional measurements, such as mass loss or CO₂ efflux.

Data from long-term N-fertilized plots (defined as +N hereafter) at an alpine site on Niwot Ridge, Colorado, provide an example of the uncertainties associated with traditional measurements. Fertilized plots in dry meadow communities of alpine tundra at this site have received annual inputs of 10 g N m⁻² yr⁻¹ since 1990 (ref. 11) (see Methods section). The decade of N fertilization has greatly increased productivity; over the past ten years, an average of 72 g m⁻² yr⁻¹ of extra C has entered the fertilized plots (Table 1). Despite these increases in productivity, soil C concentrations are not significantly different between control and +N plots (Table 1). Against the large standing pools of SOM carbon on Niwot Ridge and elsewhere, large changes in C inputs or alteration of turnover times rarely result in statistically significant changes in soil C concentrations^{12,13}.

In contrast to bulk C values, the results of radiocarbon, ¹³C and compound-specific analyses of soil C on Niwot Ridge all point to an acceleration of intermediate-age SOM turnover, and a possible increase in the stabilization of some forms of C into mineral SOM pools. Atmospheric testing of atomic weapons in the early 1950s led to a rapid increase in the ¹⁴C content of atmospheric CO₂, followed by a gradual decline to the present¹⁴ (Fig. 1). The time-dependent change in ¹⁴CO₂ can be used to examine the turnover time of intermediate-age soil C formed in equilibrium with the atmosphere. For soil samples, we performed ¹⁴C measurements on 'light' and 'heavy' SOM fractions using a high-density solution. The heavy fraction is thought to be largely recalcitrant C associated with soil minerals, with turnover times of several decades to centuries, while the more labile light fraction reflects a mixture of compounds that includes microbial biomass, partially degraded plant material and older, more humified, by-products of decomposition^{15,16}. In the dry meadow communities on Niwot Ridge, the light fraction is not significantly altered by fertilization and contributes over half of total soil C to both control and +N plots (Table 1).

The Δ¹⁴C of incoming plant C determined from summer plant harvests from 1990 to the present is similar to the Δ¹⁴C of

Table 1 Productivity and soil C and N concentrations on Niwot Ridge

	Above ground productivity (g m ⁻² yr ⁻¹)	%C	Soil C in the light fraction (%)	%N	C:N ratio
Control	151 (11)*	8.61 (1.12)*	55.12 (2.14)*	0.67 (0.07)*	12.65 (0.74)*
+N	223 (16) [†]	10.40 (1.24)*	60.44 (1.91)*	0.83 (0.08)*	12.37 (0.33)*

Values shown are means with standard errors in parentheses. Statistically significant differences below the $P < 0.05$ level are shown by contrasting symbols within columns. Average annual productivity in the control and fertilized plots is significantly different ($F = 54.54$, $P < 0.001$). +N, long term N-fertilized plots.

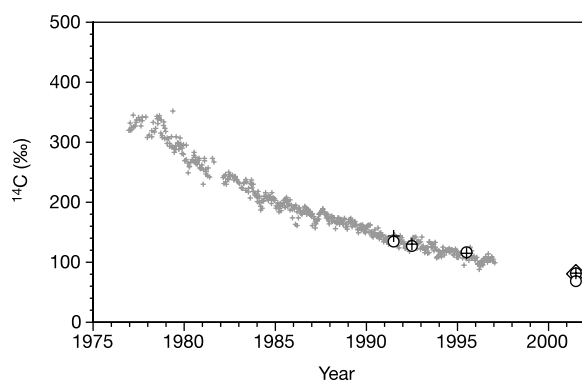


Figure 1 ^{14}C activity of Niwot Ridge plant material compared to atmospheric trends¹³. $\Delta^{14}\text{C}$ of atmospheric CO_2 from ref. 14 is shown by small grey crosses. $\Delta^{14}\text{C}$ of plant biomass for 1992, 1993, 1996 and 2001 is shown for control (open circles) and N-fertilized plots (crosses and diamond symbols).

atmospheric CO_2 , and exhibits a decline to a contemporary $\Delta^{14}\text{C}$ value of approximately 80‰ (Fig. 1). The light fraction of SOM from control plots has a $\Delta^{14}\text{C}$ of 126‰, indicative of a pool of soil C with a turnover of approximately a decade. Heavy-fraction $\Delta^{14}\text{C}$ of control plots is -13‰, reflecting C deposited mainly before atmospheric testing of nuclear weapons¹⁴. Fertilization leads to a significant reduction in light-fraction SOM $\Delta^{14}\text{C}$, to an average of 54.6‰; this value is below the 2001 input signal of approximately 80‰ (Table 2).

The most likely cause for the decline in +N light-fraction $\Delta^{14}\text{C}$ is the accelerated decomposition of highly labelled organic matter deposited roughly 10–30 years ago when atmospheric $\Delta^{14}\text{C}$ values were higher, combined with increased productivity (and inputs of less-enriched carbon) over the past decade. Changes in the structure of soil carbon after N addition suggest that relatively unaltered plant carbon resides in these soils for years to decades and then disappears as a direct or indirect result of fertilization.

Three biomarkers for a range of plant compounds that are consistent across a number of species, and reflect different plant biochemical constituents, also indicate an acceleration of the decomposition of residual plant material in fertilized soils¹⁷. Both 2-methoxy-4-vinylphenol (a biomarker for relatively un degraded plant lignin¹⁷) and two plant polysaccharide markers¹⁷ (5-methyl-2-furanone and 2-hydroxy-3-methyl-2-cyclopentenone) are substantially lower in the light fraction of fertilized plots. The concentration of the lignin biomarker is 93% lower in fertilized plots than in control plots, while the two polysaccharide markers decline by an average of 91% (Table 2). In addition, increasing lignin biomarker content in the heavy fraction of fertilized soils suggests that some plant material may be moving directly into stabilized, mineral-associated SOM pools, which may explain the small increase in the $\Delta^{14}\text{C}$ of the fertilized heavy-fraction soils.

The combination of lower light-fraction $\Delta^{14}\text{C}$ and the significant reduction in both polysaccharide and lignin plant biomarkers in fertilized soils suggests that N additions trigger an increase in the decomposition of plant-related compounds that had been resident in the soil for the past several decades. This acceleration could be mechanistically controlled by direct N effects on decomposer organisms, or indirectly by enhanced decomposition of SOM as the result of increased labile C input to soils (a 'priming' effect), or by other associated changes in plant or microbial community composition. N additions on Niwot Ridge caused a change in the dominant plant community from sedges to grasses¹¹. Previous studies of the lignin biomarkers we examined suggest that increasing grass cover should lead to greater concentration of lignin biomarkers in soil rather than to the declines observed here. The decline in both lignin and polysaccharide biomarkers provides unambiguous evidence for the acceleration of the turnover of a broad range of plant compounds in these soils^{18,19}.

$\delta^{13}\text{C}$ values in plant tissue and in lignin biomarkers provide another line of evidence for increased decomposition of the light fraction. The $\delta^{13}\text{C}$ signature of 2-methoxy-4-vinylphenol in the N-fertilized light fraction closely parallels that of unmodified plant lignin, reflecting the flush of new plant material into fertilized soils combined with the loss of older biomarker pools. In contrast, the $\delta^{13}\text{C}$ value of the same marker in control soils is enriched by 2.1‰ ($P < 0.001$, Table 2), in response to isotopic fractionation that can occur in the first stage of plant material decomposition in soils^{17,20}.

While we show several lines of evidence for N-induced acceleration of light-fraction turnover, such data cannot fully distinguish the mechanisms behind this change. However, regardless of the specific mechanism, the clear result of this study is that N additions increase the decomposition of decadal-age SOM, despite the fact that the dry meadow alpine plots analysed here are cold and dry terrestrial environments. In such environments, it is thought that decomposition is controlled primarily by temperature and perhaps by moisture²¹. Indeed, the presence of relatively unmodified plant biomarkers in alpine soils would be predicted from the standard conceptual model for decomposition. The substantial reduction in the concentration of plant biomarkers with fertilization, however, indicates that N availability exerts a major control over the decomposition of these materials in this setting. However, one important issue to resolve is whether the changes in soil C cycling resulting from high rates of fertilizer addition over short time periods are similar to the impacts caused by relatively low, but chronic, additions of N to ecosystems from atmospheric deposition.

Finally, despite clear evidence for significant changes in soil C processing, the net effects of increased N on soil C storage are not certain. Our data suggest that the responses of alpine ecosystems to fertilization include both increased productivity and increased decomposition of the light fraction of SOM, resulting in no statistically detectable change in total SOM carbon. There is a well developed literature on plant responses to increasing N inputs, the bulk of which suggests that initial increases in productivity will eventually plateau, and perhaps even decline at high levels of N

Table 2 Isotope and compound-specific data from Niwot Ridge plots

	$\Delta^{14}\text{C}$	$\delta^{13}\text{C}$	Lignin marker area (Vs)	Lignin marker $\delta^{13}\text{C}$	Polysaccharide marker area (Vs)	Polysaccharide marker $\delta^{13}\text{C}$
Control light fraction	126.12 (13.29)*	-25.60 (0.08)*	37.54 (7.37)*	-27.34 (0.41)*,†	34.11 (6.28)*	-19.44 (0.44)*
+N light fraction	54.60 (15.31)†	-26.01 (0.17)†	0.68 (0.12)†	-29.40 (0.42)*	1.08 (0.07)†	-19.76 (0.55)*,†
Control heavy fraction	-13.31 (7.91)†	-24.78 (0.07)†	2.10 (1.01)†	-25.90 (0.73)†	0.95 (0.33)†	-21.08 (0.18)†
+N heavy fraction	-1.05 (7.37)†	-24.90 (0.07)†	2.56 (0.29)†	-25.35 (0.21)†	2.80 (0.41)†	-20.42 (0.17)*,†

Values shown are means with standard errors in parentheses (see Supplementary Information for individual $\Delta^{14}\text{C}$ values). Contrasting symbols within each column indicate significant differences determined from Tukey post-hoc analyses following two-way ANOVA. Statistical results for $\Delta^{14}\text{C}$ illustrate significant fertilization ($F = 76.45$, $P < 0.001$), SOM density ($F = 3.23$, $P < 0.01$) and density by fertilization interaction ($F = 13.59$, $P < 0.001$) effects. Soil $\delta^{13}\text{C}$ is affected by SOM density ($F = 82.49$, $P < 0.001$) and fertilization ($F = 6.50$, $P < 0.015$). The concentrations of 2-methoxy-4-vinylphenol, a plant lignin biomarker, are shown as lignin marker area (Vs is normalized abundance). Significant effects are shown for fertilization ($F = 18.19$, $P < 0.001$), SOM density ($F = 18.19$, $P < 0.001$) and density by fertilization interactions ($F = 21.40$, $P < 0.001$). The $\delta^{13}\text{C}$ of the marker for lignin shows significant density ($F = 30.70$, $P < 0.001$) and density $\times$ treatment interactions ($F = 2.32$, $P < 0.022$). Variations in plant polysaccharides were analysed by combining normalized abundance (to 1 mg sample) values for 5-methyl-2-furanone and 2-hydroxy-3-methyl-2-cyclopentenone and these values show significant fertilization ($F = 24.40$, $P < 0.001$), SOM density ($F = 30.04$, $P < 0.001$) and density by fertilization interactions ($F = 24.00$, $P < 0.001$). The $\delta^{13}\text{C}$ of the polysaccharide markers is influenced by SOM density ($F = 8.758$, $P < 0.001$). Vs, volt second.

loading²². Despite the fact that soil C pools greatly exceed vegetation storage, there is no comparable conceptual or quantitative model for soil C responses to changing N availability. An improved understanding of the mechanisms of soil C response to changing N availability, and more accurate SOM carbon measurement techniques will be needed for accurate modelling^{5,21} of ecosystem C storage in response to a changing N cycle. □

Methods

Site description

The experimental plots were located at the National Science Foundation, Long-Term Ecological Research site on Niwot Ridge, an 8-km ridge that extends east from the continental divide in the Front Range of the Colorado Rocky Mountains (<http://culter.colorado.edu>). The soils of the dry meadow community are Inceptisols with texture of 39% sand, 28% silt and 23% clay²³. The N-fertilization experiment was carried out in a dry meadow community dominated by *Kobresia myosuroides* and including other species in the Cyperaceae. Forbs constitute the other major growth form. Experimental fertilization plots were established in dry meadow sites in May 1990, and consist of five replicate 2 × 2 m plots including control and fertilizer addition plots where N has been applied every year since 1990 at an average rate of 25 g N m⁻² yr⁻¹ as urea-N for the first two years and at 10 g N m⁻² yr⁻¹ thereafter. For reference, atmospheric deposition rates at Niwot Ridge are 0.5 g N m⁻² yr⁻¹ (ref. 24). Full details on the plot design, fertilizer additions and plant measurements can be found in ref. 11.

Soil measurements

We collected soils from the top 10 cm of the experimental plots. For most analyses, we sieved soils (2 mm) and then separated the soils into light and heavy fractions by floating soils in sodium polytungstate (density ~1.6 g cm⁻³)¹⁵. For C concentrations, ¹⁴C/¹²C and ¹³C/¹²C determinations, we took two samples from each plot and averaged these for a single plot-value used in statistical tests. Targets for ¹⁴C measurement were prepared at the Laboratory for Radiocarbon Preparation and Research at the Institute for Arctic and Alpine Research and analysed at the National Ocean Sciences Accelerator Mass Spectrometer Facility at the Woods Hole Oceanographic Institution. $\Delta^{14}\text{C}$ values were decay corrected to the date of collection²⁵ and presented as per mil deviations from fraction modern (F_m) = 1, where $F_m = (^{14}\text{C}/^{12}\text{C})_{\text{sample}} / (^{14}\text{C}/^{12}\text{C})_{\text{on}}$ and the subscript on denotes the sample value normalized to a $\delta^{13}\text{C}$ value of -25‰ and the subscript on indicates the radiocarbon standard value normalized to a $\delta^{13}\text{C}$ value of -25‰ (ref. 24) (see Supplementary Information for individual radiocarbon results).

Compound-specific analyses

Compound-specific analyses were performed at the Max Planck Institute for Biogeochemistry in Jena, Germany, using pyrolysis-gas chromatography/mass spectrometry-combustion interface-isotope ratio-mass spectrometry (py-GC/MS-C-IRMS). Briefly, samples were pyrolysed for 9.9 s in a 0316 Fisher pyrolyser using a ferromagnetic tube with a Curie temperature of 590 °C. Pyrolysis products were separated on a BPX 5 column (60 m × 0.32 mm, film thickness 1.0 μm) and detected on a Thermoquest GCQ operated at 70 eV. Compound-specific ¹³C measurements were made on a Finnigan MAT Delta^{plus}XL. Additional analytical details are available in ref. 17.

Statistical calculations

Our statistical calculations were carried out with a two-way factorial analysis of variance (ANOVA) combined with Tukey post-hoc pair-wise comparisons to examine statistical differences between soil densities (light versus heavy) and control versus fertilization treatments.

Received 29 April; accepted 12 September 2002; doi:10.1038/nature01136.

- Post, W. M., Emanuel, W. R., Zinke, P. J. & Stangenberger, A. G. Soil carbon pools and world life zones. *Nature* **298**, 156–159 (1982).
- Trumbore, S. E., Chadwick, O. A. & Amundson, R. Rapid exchange between soil carbon and atmospheric carbon dioxide driven by temperature change. *Science* **272**, 393–396 (1996).
- Epstein, H. E., Burke, I. C. & Lauenroth, W. K. Regional patterns of decomposition and primary production rates in the US Great Plains. *Ecology* **83**(2), 320–327 (2002).
- Vitousek, P. M. *et al.* Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* **7**(3), 737–750 (1997).
- Townsend, A. R., Braswell, B. H., Holland, E. A. & Penner, J. E. Spatial and temporal patterns in terrestrial carbon storage due to deposition of fossil fuel nitrogen. *Ecol. Appl.* **6**(3), 806–814 (1996).
- Vitousek, P. M. Nutrient cycling and nutrient use efficiency. *Am. Nat.* **119**, 553–572 (1982).
- Melillo, J. M. & Aber, J. D. Nitrogen and lignin control of hardwood leaf litter decomposition dynamics. *Ecology* **63**(3), 621–626 (1982).
- Hobbie, S. E. Interactions between litter lignin and soil nitrogen availability during leaf litter decomposition in a Hawaiian montane forest. *Ecosystems* **3**, 484–494 (2000).
- Resh, S. C., Binkley, D. & Parrotta, J. A. Greater soil carbon sequestration under nitrogen-fixing trees compared with *Eucalyptus* species. *Ecosystems* **5**, 217–231 (2002).
- Berg, B. & Matzner, E. Effect of N deposition on decomposition of plant litter and soil organic matter in forest systems. *Environ. Rev.* **5**, 1–25 (1997).
- Bowman, W. D., Theodose, T. A., Schardt, J. C. & Conant, R. T. Constraints of nutrient availability on primary production in two alpine communities. *Ecology* **74**, 2085–2098 (1993).
- Schlesinger, W. H. & Lichter, J. Limited carbon storage in soil and litter of experimental forest plots under increased CO₂. *Nature* **411**, 466–468 (2001).

- Magill, A. H. *et al.* Biogeochemical response of forest ecosystems to stimulated chronic nitrogen deposition. *Ecol. Appl.* **7**(2), 402–415 (1997).
- Levin, I. & Kromer, B. *Trends: A Compendium of Data on Global Change* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US DOE, Oak Ridge, TN, 1997).
- Trumbore, S. E. Comparison of carbon dynamics in tropical and temperate soils using radiocarbon measurements. *Glob. Biogeochem. Cycles* **7**(2), 275–290 (1993).
- Gaudinski, J. B., Trumbore, S. E., Davidson, E. A. & Zheng, S. Soil carbon cycling in a temperate forest: radiocarbon-based estimates of residence times, sequestration rates and partitioning of fluxes. *Biogeochemistry* **51**(1), 33–69 (2000).
- Gleixner, G., Poirier, N., Bol, R. & Balesdent, J. Molecular dynamics of organic matter in a cultivated soil. *Org. Geochem.* **33**(3), 357–366 (2002).
- Nierop, K. G. J., Pulleman, M. M. & Marinissen, J. C. Y. Management induced organic matter differentiation in grassland and arable soil: a study using pyrolysis techniques. *Soil Biol. Biochem.* **33**, 755–764 (2001).
- Kuder, T. & Kruege, M. A. Preservation of biomolecules in sub-fossil plants from raised peat bogs—a potential paleoenvironmental proxy. *Org. Geochem.* **29**(5–7), 1355–1368 (1998).
- Schmidt, H. L. & Gleixner, G. *Stable Isotopes. Integration of Biological, Ecological and Geochemical Processes* (ed. Griffith, H.) 13–25 (BIOS, Oxford, UK, 1998).
- Schimmel, D. S. *et al.* Climatic, edaphic and biotic controls over storage and turnover of carbon in soils. *Glob. Biogeochem. Cycles* **8**(3), 279–293 (1994).
- Aber, J. *et al.* Nitrogen saturation in temperate forest ecosystems—hypotheses revisited. *Bioscience* **48**(11), 921–934 (1998).
- Burns, S. F. *Alpine Soil Distribution and Development, Indian Peaks, Colorado Front Range*. PhD Univ. Colorado (1980).
- Williams, M. W., Hood, E. & Caine, N. Role of organic nitrogen in the nitrogen cycle of a high elevation catchment, Colorado Front Range. *Wat. Resour. Res.* **37**(10), 2569–2581 (2001).
- Stuiver, M. & Polach, H. A. Discussion: Reporting of ¹⁴C data. *Radiocarbon* **19**(3), 355–363 (1977).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank D. Schimmel, S. Hobbie, J. Harden, T. Seastedt and C. Cleveland for comments on an earlier draft of this manuscript. This work was supported by the Andrew Mellon Foundation, the NSF-LTER program, the US Geological Survey and the Max Planck Institute for Biogeochemistry.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J. N. (e-mail: jneff@usgs.gov).

Quantifying nitrogen-fixation in feather moss carpets of boreal forests

Thomas H. DeLuca^{*†}, Olle Zackrisson[†], Marie-Charlotte Nilsson[†] & Anita Sellstedt[‡]

^{*} School of Forestry, The University of Montana, Missoula, Montana 59812, USA
[†] Department of Forest Vegetation Ecology, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden
[‡] Umeå Plant Science Center, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

Biological nitrogen (N) fixation is the primary source of N within natural ecosystems¹, yet the origin of boreal forest N has remained elusive. The boreal forests of Eurasia and North America lack any significant, widespread symbiotic N-fixing plants^{1–6}. With the exception of scattered stands of alder in early primary successional forests⁷, N-fixation in boreal forests is considered to be extremely limited. Nitrogen-fixation in northern European boreal forests has been estimated² at only 0.5 kg N ha⁻¹ yr⁻¹; however, organic N is accumulated in these ecosystems at a rate of 3 kg N ha⁻¹ yr⁻¹ (ref. 8). Our limited understanding of the origin of boreal N is unacceptable given the extent of the boreal forest region, but predictable given our imperfect knowledge of N-fixation^{1,9}. Herein we report on a N-fixing symbiosis between a cyanobacterium (*Nostoc* sp.) and the

ubiquitous feather moss, *Pleurozium schreberi* (Bird) Mitt. that alone fixes between 1.5 and 2.0 kg N ha⁻¹ yr⁻¹ in mid- to late-successional forests of northern Scandinavia and Finland. Previous efforts have probably underestimated N-fixation potential in boreal forests.

To document the occurrence and distribution of N-fixation in *P. schreberi*, we collected moss samples from 27 field sites in northern Sweden, Norway and Finland between latitudes 62–70° N and between May and November 2001. Nitrogen-fixation rates associated with the moss were estimated by measuring acetylene reduction rates in field incubations. Acetylene reduction rates were found to be linear with time over 24 h, so we performed 24-h field incubations to capture diurnal influences on N-fixation rates. We found substantial levels of N-fixation in *P. schreberi* at 23 of the 27 sites (Fig. 1). The average acetylene reduction rate for these 27 sites was 88 $\mu\text{mol m}^{-2} \text{d}^{-1}$ and averages for individual sites ranged from levels below the detection limit to 460 $\mu\text{mol m}^{-2} \text{d}^{-1}$. Acetylene reduction rates were calibrated using $^{15}\text{N}_2(\text{g})$ and found to equate to a ratio of 3 moles of acetylene reduced for each mole of N_2 . From these 27 sites we might estimate average N-fixation rates at 1.6 kg N ha⁻¹ yr⁻¹ (assuming 65% ground cover for *P. schreberi*¹⁰ and a 200-d growing season).

We also conducted a season-long *in situ* field assessment of N-fixation rates at the Reivo forest reserve, northern Sweden. Nitrogen-fixation rates were monitored from just before snow melt in early May to first snow accumulation in early December (approximately 210 d of field observation). Nitrogen-fixation activity, again measured by acetylene reduction, was initially low under snow pack ($\sim 10 \mu\text{mol m}^{-2} \text{d}^{-1}$), but rates climbed rapidly in June to 250 $\mu\text{mol m}^{-2} \text{d}^{-1}$ (Fig. 2). Average acetylene reduction rates reached a seasonal peak of 370 $\mu\text{mol m}^{-2} \text{d}^{-1}$ in September following a brief drop during mid-summer (Fig. 2). The highest individual observation of acetylene reduction was 4,265 $\mu\text{mol m}^{-2} \text{d}^{-1}$ on 30 August 2001. Nitrogen-fixation rates dropped again in October with the first hard frost, but then recovered to average acetylene reduction rates of 170 $\mu\text{mol m}^{-2} \text{d}^{-1}$ through mid-September and remained near 100 $\mu\text{mol m}^{-2} \text{d}^{-1}$ through early November in spite of low (subzero) nighttime temperatures. The average rate of acetylene reduction for the 210-d period was 112 $\mu\text{mol m}^{-2} \text{d}^{-1}$ which would equate to an N-fixation rate of 2.1 kg N ha⁻¹ yr⁻¹ (given 65% ground cover for *P. schreberi* and a molar ratio of 3:1 acetylene reduced to N_2 reduced). By using a time-weighted average for each sample point, however, we arrive at a more conservative, and perhaps more accurate estimate of 1.7 kg N ha⁻¹ yr⁻¹.

An N-fixation rate of 1.7 kg N ha⁻¹ yr⁻¹ in *P. schreberi* alone is

approximately ten times greater than previous estimates for the total moss bottom layer in northern European boreal forests² and three times greater than whole watershed estimates of N-fixation in these forests including sphagnum mires and N-fixing woody plants². Cleveland *et al.*¹ have estimated total N-fixation across North American and Eurasian boreal forests and woodlands at 1.5–2 kg N ha⁻¹. Adding the N contribution of *P. schreberi* to this total value would potentially double their estimate of total N fixation in these forests.

Microscopic observation of the moss tissue confirmed the presence of an epiphytic *Nostoc* sp. (*Nostoc* cf. *sphaericum* Vaucher ex Bornet & Flahault) within the incurve of the moss leaf (Fig. 3). Both hormogonia (a motile and colonization stage common in symbiotic *Nostoc* spp.¹¹) and coiled, heterocyst N-fixation stages were observed in moss tissues. The presence of *Nostoc* within the leaf incurve conceals the cyanobacteria from microscopic observation of moss leaf surfaces (Fig. 3) and may explain why epiphytic cyanobacteria were not previously observed in *P. schreberi*¹².

Nitrogen-fixing cyanobacteria and heterotrophic bacteria have been reported to be capable of intercellular, epiphytic and free-living growth associated with certain species of moss in Antarctic, Arctic, grassland and boreal settings^{12–17}. High rates of N-fixation in sphagnum mosses (by epiphytic cyanobacteria) were reported in Swedish mires¹⁷, however, feather moss samples collected within the same region were found to have low or non-measurable rates of N-fixation². All samples collected by these researchers² were taken in July or August, a potential dormant season for feather mosses¹⁸ and a period in which we observed reduced fixation rates (Fig. 2). In a study of N-fixation rates in lichen and mosses in the Alaskan boreal forest, Alexander and Billington¹² reported their highest field observation for N-fixation (2,200 $\mu\text{mol m}^{-2} \text{d}^{-1}$) to be associated with *P. schreberi*, but their attempts to observe epiphytic cyanobacteria failed and so this value went essentially unnoticed by the scientific community.

Nitrogen-fixation in *P. schreberi* has significant regional and global implications. *P. schreberi* is the most common moss on Earth, being found throughout North and South America, Greenland, Asia, Europe and Africa and ranging from sea level (just off active beaches) to 5,000 m elevation¹⁹. In addition, *P. schreberi* is the most prevalent feather moss in the boreal forest region (which accounts for 14% of the total global land mass). Feather moss cover increases with successional age¹⁰, ultimately forming a continuous carpet (dominated by *P. schreberi*)²⁰ and accounting for 60 to over

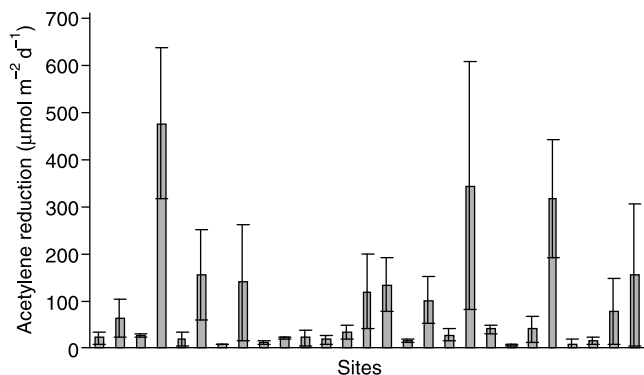


Figure 1 Average acetylene reduction rates ($\mu\text{mol m}^{-2} \text{d}^{-1}$) for *P. schreberi* at 27 different boreal sites in northern Europe between latitudes of 62–70° N and longitudes of 13–20° E. Each bar indicates a distinct site and the vertical bars represent ± 1 s.e., $n = 4$ –12.

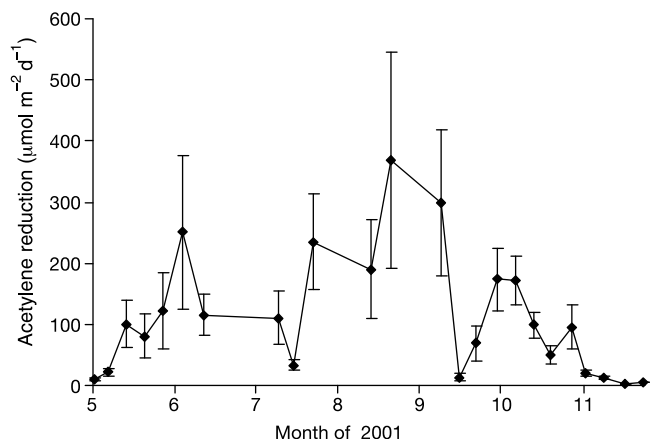


Figure 2 Average acetylene reduction rates ($\mu\text{mol m}^{-2} \text{d}^{-1}$) for *P. schreberi* from May to early December 2001 at the Reivo forest reserve in northern Sweden. Each point represents the mean of one sampling period (1–3 weeks apart). Vertical bars represent ± 1 s.e., $n = 24$.

80% of ground cover in boreal forests^{10,12,20–22}. These dense feather moss carpets have been found to account for greater primary productivity than the overstory tree species in Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), and black spruce (*Picea mariana*) boreal forests^{12,20,22}. Carpets of feather moss have been described as a “sponge” for nutrients captured from canopy throughfall²² and are often considered to reduce forest productivity owing to forest nutrient retention. Soil scarification and herbicides are commonly used in the boreal forest to reduce site competition by feather mosses and enhance growth rates of planted merchantable tree species.

It is well established that N-fixation is a major driver of ecosystem processes^{23,24}. Plant community dynamics can be influenced profoundly by the presence of N-fixing organisms. It is typically thought, however, that N-fixation is only important in early primary succession²⁵ and that N accumulation precludes the need for the energy consumptive process of N-fixation. Although the organic N pool in the boreal forest is large, very little N is available for plant or microbial use⁴. This lack of available N may create a demand for N-fixation in mid- to late succession, but it is likely that N transfer from feather moss carpets to trees or understory shrubs is tortuous²⁶. Regardless of transfer rates, our results suggest that the *P. schreberi*–*Nostoc* symbiosis is an important contributor to N accumulation and N cycling in boreal forests as well as in tundra ecosystems where *P. schreberi* is also an important moss species¹⁹.

The findings presented here clearly identify *P. schreberi* as a host for an associative or symbiotic *Nostoc* sp. The spatial and temporal variability in N-fixation rates associated with *P. schreberi* and the unusual seasonal pattern (greatest fixation rates in late summer to early fall) may explain why this association has been greatly overlooked in the past. The observed rates of N-fixation coupled with the ubiquitous, dense nature of *P. schreberi*-dominated feather moss carpets produce a significant N contribution to the total boreal N

balance that must be accounted for in the modelling and management of boreal forest ecosystems. □

Methods

Replicate field samples were collected at 27 sites in northern Sweden, Norway and Finland between latitudes of 62–70° N and longitudes of 13–20° E. Nitrogen-fixation rates were estimated by use of the acetylene reduction method²⁷. Samples of ten moss shoots (approximately 2.8 cm² of moss or 0.10–0.35 g dry weight basis) were placed into 20-ml culture tubes and sealed with a rubber septum. A 1.8-ml portion of air was removed from the headspace using a syringe and replaced with 1.8 ml of reagent grade acetylene. The tubes were then nestled back into the moss layer and allowed to incubate under natural conditions over a 24-h period. A 1-ml sample of the headspace was then removed using a gas-tight syringe and analysed for ethylene concentration using a Schimadzu gas chromatograph equipped with a flame ionization detector and Porapak Q column. Nitrogen-fixation activity was recorded in units of $\mu\text{mol ethylene cm}^{-2} \text{ h}^{-1}$. Acetylene reduction rates were tested for linearity at both 7 °C and 22 °C by analysing ethylene production in tubes with moss at 0.08, 0.25, 0.5, 1, 2, 4, 8, 16 and 24 h. Acetylene reduction was found to be linear through the 24-h period at both temperatures. Localized areas of high fixation rates were identified at most sites. We believe these areas of high fixation to be important contributions to overall N-fixation rates at individual sites. Results are therefore reported as arithmetic mean values rather than as median values to avoid elimination of the important contribution of localized areas of rapid N-fixation.

An assessment of season-long N-fixation rates was conducted in a late-succession open Scots pine forest at the Reivo forest reserve in northern Sweden (65° 46' 65" N, 19° 05' 71" E). Moss samples were collected under snow pack in May and then collected every 1–3 weeks throughout the summer and autumn until snow accumulation in December. Samples were collected from 24 separate 1 × 1 m field plots within an area of 5 ha on homogeneous understory vegetation wherein *P. schreberi* was dominant (65% cover) in the bottom layer and *Vaccinium* spp. and *Empetrum hermaphroditum* were dominant in the field layer. The percentage ground cover of field and bottom layer vegetation was estimated at the species level in 12 replicated large plots (11.3 m diameter)²⁸. Tubes were field incubated using the techniques described above and temperature and light were continuously monitored at two of the field plots over the 24-h incubation period. Samples were collected and field incubated on a total of 24 occasions from 16 May 2001 to 3 December 2001.

Moss samples were analysed for the presence of N-fixing cyanobacteria and heterotrophic bacteria. Leaves of moss were examined at × 200–400 under a fluorescence microscope with a green filter, for the presence of cyanobacteria. Leaves from two positions on moss stems were also removed and placed into culture tubes containing 10 ml of sterile distilled water. Samples were shaken and 1 ml of the suspension was transferred to plates or culture tubes containing BG11 N-free media. Inoculated media was allowed to incubate at 25 °C in full light for 6 weeks and observed for the presence of cyanobacteria at × 400.

To allow conversion of acetylene reduction values to N-fixation rates, the acetylene reduction method was calibrated by incubating moss samples with acetylene and with ¹⁵N₂ gas²⁹. Five replicate samples of ten moss shoots each were placed into 20-ml tubes as described above and sealed with a rubber septum. Samples were tested for acetylene reduction over a 1-h period as described above, following which the septa were removed for several hours to allow all acetylene to escape. The septa were replaced and 20% of the headspace was removed using a gas-tight syringe and replaced with N₂ gas enriched at 98% ¹⁵N₂. Moss samples were then incubated for one hour at 50% full light and at 22 °C. Control tubes containing moss with no ¹⁵N₂ gas and control tubes containing no moss were incubated in parallel with the treated ¹⁵N₂ moss. Moss samples were dried at 60 °C, ground in a ball mill and analysed for ¹⁵N enrichment using a continuous-flow isotopes ratio mass spectrometer. Rates of ¹⁵N₂ fixation equate to a molar ratio of 3 moles of acetylene reduced to 1 mole of N₂.

Acetylene reduction rates in 20-ml glass tubes were converted to aerial N-fixation rates by assuming a 65% ground cover of *P. schreberi*¹⁰, a molar ratio of 3 moles acetylene reduced to one mole N₂ reduced, and an average of 2.8 cm² for ten shoots of moss. The length of period for fixation was directly measured in the field experiment at Reivo and was assumed to be 200 d for the 27 sites across northern Scandinavia and Finland.

Received 19 April; accepted 15 July 2002; doi:10.1038/nature01051.

- Cleveland, C. C. *et al.* Global patterns of terrestrial biological nitrogen (N₂) fixation. *Glob. Biogeochem. Cycles* **13**, 623–645 (1999).
- Rosen, K. & Lindberg, T. Biological nitrogen-fixation in coniferous forest watershed areas in central Sweden. *Holarct. Ecol.* **3**, 137–140 (1980).
- Walter, H. & Breckle, S.-W. *Ecological Systems of the Geobiosphere. Temperate and Polar Zonobiomes of Northern Eurasia* (Springer, Berlin, 1989).
- Tamm, C. O. *Nitrogen in Terrestrial Ecosystems* 34–49 (Springer, Berlin, 1991).
- Longton, R. E. *Bryophytes and Lichens in an Changing Environment* (eds Bates, J. W. & Farmer, A. M.) 32–76 (Clarendon, Oxford, UK, 1992).
- Shugart, R., Leemans, R. & Bonan, G. B. *A System Analysis of Global Boreal Forest* 565 (Cambridge Univ. Press, Cambridge, UK, 1992).
- Vogel, J. G. & Gower, S. T. Carbon and nitrogen dynamics of boreal jack pine stands with and without a green alder understory. *Ecosystems* **1**, 386–400 (1998).
- Wardle, D. A., Zackrisson, O., Hörnberg, G. & Gallet, C. The influence on island area on ecosystem properties. *Science* **277**, 1296–1299 (1997).
- Binkley, D., Son, Y. & Valentine, D. W. Do forests receive occult inputs of nitrogen? *Ecosystems* **3**, 321–331 (2000).
- DeLuca, T. H., Nilsson, M.-C. & Zackrisson, O. Nitrogen mineralization and phenol accumulation along a fire chronosequence in northern Sweden. *Oecologia* advance online publication, 28 August 2002 (DOI 10.1007/S00442-002-1025-2).

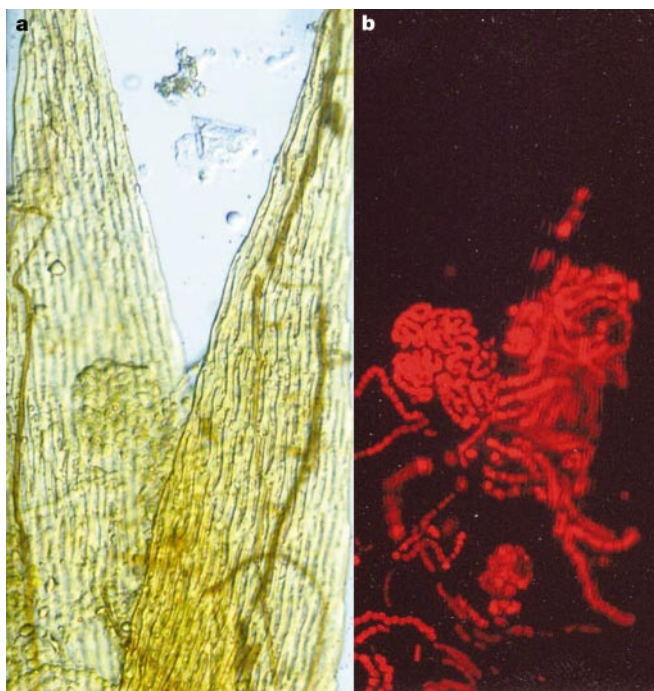


Figure 3 A micrograph of a section of moss leaf present at × 200 magnification. **a**, Under light microscope; and **b**, under ultraviolet-fluorescence micrograph with a green filter. Coiled chains of *Nostoc* are hidden in the leaf under light microscopy, but are readily observed as the red cells under ultraviolet-fluorescence microscopy with a green filter. (The photographs were taken by P. Lundgren and U. Rasmussen.)

11. Rai, A. N., Soderback, E. & Bergman, B. Cyanobacteria-plant symbioses. *New Phytol.* **147**, 449–481 (2000).
12. Alexander, V. & Billington, M. M. *Forest Ecosystems in the Alaskan Taiga: a Synthesis of Structure and Function* (eds Van Cleve, K., Chapin, F. S., Flanagan, P. W., Viereck, L. A. & Dyrness, C. T.) 112–120 (Springer, New York, 1986).
13. Christie, P. Nitrogen fixation in two contrasting Antarctic bryophyte communities. *J. Ecol.* **75**, 73–93 (1987).
14. Reddy, G. B. & Giddens, J. Nitrogen fixation by moss-algal association in grasslands. *Soil Biol. Biochem.* **13**, 537–538 (1981).
15. Chapin, D. M., Bliss, L. C. & Bledsoe, L. J. Environmental regulation of nitrogen fixation in a high arctic lowland ecosystem. *Can. J. Bot.* **69**, 2744–2755 (1991).
16. Pandey, K. D., Kashy, A. K. & Gupta, R. K. Nitrogen fixation by cyanobacteria associated with moss communities in Schirmacher Oasis, Antarctica. *Israel J. Bot.* **41**, 187–198 (1992).
17. Granhall, U. & Selander, H. Nitrogen fixation in a subarctic mire. *Oikos* **24**, 8–15 (1973).
18. Whitehead, D. & Gower, S. T. Photosynthesis and light-use efficiency by plants in a Canadian boreal forest ecosystem. *Tree Physiol.* **21**, 925–929 (2001).
19. Kuc, M. The northernmost extension of the moss *Pleurozium schreberi* (Brid.) Mitt. in the Canadian High Arctic. *Can. Field Nat.* **111**, 630–633 (1997).
20. Foster, D. R. Vegetation development following fire in *Picea mariana* (black spruce)–*Pleurozium* forests of south-eastern Labrador, Canada. *J. Ecol.* **73**, 517–534 (1985).
21. Melkønen, E. Annual primary production and nutrient cycle in some Scots pine stands. *Commun. Inst. Forest. Fenn.* **84**, 1–76 (1974).
22. Oechel, W. C. & Van Cleve, K. *Forest Ecosystems in the Alaskan Taiga: A Synthesis of Structure and Function* (eds Van Cleve, K., Chapin, F. S., Flanagan, P. W., Viereck, L. A. & Dyrness, C. T.) 121–137 (Springer, New York, 1986).
23. Vitousek, P. M. & Hobbie, S. Heterotrophic nitrogen fixation in decomposing litter: Patterns and regulation. *Ecology* **81**, 2366–2376 (2000).
24. Chapin, F. S. III, Vitousek, P. M. & Van Cleve, K. The nature of nutrient limitation in plant communities. *Am. Nat.* **127**, 48–58 (1986).
25. Rastetter, E. B. *et al.* Resource optimization and symbiotic nitrogen fixation. *Ecosystems* **4**, 369–388 (2001).
26. Van Cleve, K. & Alexander, V. in *Terrestrial Nitrogen Cycles* Ecol. Bull. 33 (eds Clark, F. E. & Rossell, T.) 375–404, (Swedish Natural Science Research Council, Stockholm, 1981).
27. Mattsson, U. & Sellstedt, A. Hydrogenase in Frankia KB5: Expression of and relation to nitrogenase. *Can. J. Microbiol.* **46**, 1091–1095 (2000).
28. Hägglund, B. & Lundmark, J. E. Site index estimation by means of site properties. *Studia Forestalia Suecica* **138**, 1–34 (1977).
29. Weaver, R. W. & Danso, S. K. A. *Methods of Soil Analysis Part 2 Biochemical and Microbiological Properties* (eds Weaver, R., Angle, J. S. & Bottomley, P. S.) 985–1018 (Soil Science Society of America, Madison, WI, 1994).

Acknowledgements We thank R. Sheridan, B. Bergman, U. Rasmussen, and J. Johansen for their advice and assistance.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.H.D. (e-mail: thd@forestry.umn.edu).

Honeybee colonies achieve fitness through dancing

Gavin Sherman & P. Kirk Visscher

Department of Entomology, University of California, Riverside, California 92521, USA

The honeybee dance language, in which foragers perform dances containing information about the distance and direction to food sources, is the quintessential example of symbolic communication in non-primates^{1,2}. The dance language has been the subject of controversy^{3,4}, and of extensive research into the mechanisms of acquiring^{5,6}, decoding^{7,8} and evaluating⁹ the information in the dance. The dance language has been hypothesized, but not shown, to increase colony food collection^{1,9,10}. Here we show that colonies with disoriented dances (lacking direction information) recruit less effectively to syrup feeders than do colonies with oriented dances. For colonies foraging at natural sources, the direction information sometimes increases food collected, but at other times it makes no difference. The food-location information in the dance is presumably important

when food sources are hard to find, variable in richness and ephemeral. Recruitment based simply on arousal of foragers and communication of floral odour, as occurs in honeybees¹, bumble bees¹¹ and some stingless bees¹², can be equally effective under other circumstances. Clarifying the condition-dependent payoffs of the dance language provides new insight into its function in honeybee ecology.

The dance resembles a miniaturized re-enactment of the flight to the food source. Elements of the dance (repeated waggle runs) encode the direction and distance to the food source. The direction of the waggle run corresponds to the angle of the heading to the food source relative to the current sun azimuth, and is performed relative to a sensory reference. In *Apis mellifera* the dances are usually done on a vertical comb in a dark nest cavity, and the reference is upward. In *Apis florea* and in some circumstances in other *Apis* species, the dance is performed on a horizontal surface with a reference to the sun azimuth, detected from a direct view of the sun, sky polarization patterns or landmarks¹³. When on horizontal comb in the dark or in diffused light, bees still dance, but in random directions, and the indication of distance is also disrupted¹. Bees will interpret spots of artificial light as the Sun or as particular regions of the polarized blue sky^{1,5,14}. In addition to distance and direction information, bees that follow dances learn the odour of the food, and that a source rich enough to merit recruitment exists⁹; such richness and odour information also occur in the dances of other social bee species¹¹. After honeybee dances were first decoded by von Frisch¹, controversy arose about whether bees that follow dances actually decode the distance and direction information, or instead rely exclusively on odour^{3,4}. Subsequent experiments established that bees can decode the dances^{4,7}, but the relative role of odour and vector information has been little studied¹⁵. Mechanistic and descriptive studies suggest that the dance makes the nest an information centre where communication allows a colony to direct most of its foragers to the richest sources found, increasing the colony food collection^{1,4,10}, but this hypothesis has not been effectively tested¹⁶.

To test the effect of the dance language, we established a diffuse-light treatment in which bees performed completely disoriented dances, and an oriented-light treatment in which bees performed well-oriented dances. Former studies comparing disoriented and oriented dances^{1,3,15} compared vertical and horizontal hives, which may differ in factors other than dance orientation. We used two-frame horizontal observation hives that could be illuminated with either treatment: light from three 25-W bulbs diffused through

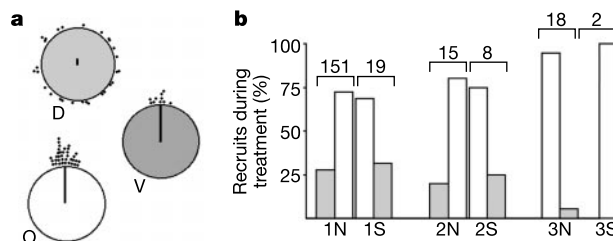


Figure 1 Disoriented and oriented dances and recruitment. **a**, Variation in direction of waggle runs within single dances of dancers in colonies with diffuse-light treatment (D), oriented-light treatment (O) and a vertical colony (V). Each point on the circumference represents one waggle run, and the radial line the direction and length of the mean vector. The dances shown are the ones closest to the means of mean vector length for ten recordings of dances of each treatment, rotated to align mean vectors, which indicate arbitrary directions of natural food sources. **b**, Percentage of total recruits per colony per trial (totals above bars) from each of a pair of colonies (N and S) that arrived while receiving oriented-light (open bars) and diffuse-light (shaded bars) treatments in three trials. The reciprocal design ensures that these changes were not due to changes in other available forage during the day. Wind was consistent throughout the experiment, so wind changes could not account for the result²⁰.

white translucent Plexiglas suspended above the colony, or unidirectional light from one 75-W bulb above a sheet of transparent Plexiglas plus a fluorescent 'black light' providing short wavelengths. The directions of waggle runs transcribed from videotapes of single dances in the diffuse-light treatment did not differ from a circular random distribution (mean vector length $\pm$ s.e. in ten dances averaged 0.198 ± 0.038 , Rayleigh test (ref. 17), $P > 0.5$ for nine of the ten), as shown in Fig. 1a. The dances in the oriented-light treatment were highly directionally oriented (mean vector length averaged 0.950 ± 0.012 , $P < 0.001$ for each of ten dances), similar to dances on vertical comb (where mean vector length averaged 0.984 ± 0.004).

To verify that these light treatments affected recruitment, we compared the recruitment to feeders when colonies had disoriented and oriented dances in reciprocal-treatment experiments. This design provided tests for effects of light treatment, colony and time of day, as well as controlling for wind and the possibility of locale-specific odours^{3,4}. An equal number of foragers were trained from each colony to a feeder, and individually tagged for identification. Each colony's feeder was in a different direction. The colonies received opposite light treatments in the first half of the observation; then the treatments were switched. Unmarked recruits arriving at each feeder were captured and recorded. In three trials at 250, 400 and 460 m from the colony, significantly more recruits arrived at feeders while the colony had oriented dances than arrived when it had disoriented dances (trial 1: χ^2 test for homogeneity, $P < 0.001$, trials 2 and 3: Fisher exact test, $P < 0.02$). There were no time-of-day effects on recruit numbers (same tests, $P > 0.7$); there was a colony effect in trial 1 ($P < 0.001$) but not in trials 2 and 3 ($P > 0.25$ and 0.15 , respectively). To test whether foragers were less likely to dance when unable to do oriented dances, we made counts of the number of tagged bees dancing in each colony, under both light treatments. An analysis of variance (ANOVA) of this data (transformed as square root (count + 0.5) to improve normality and homoscedasticity) showed no significant effect of light treatment on the mean number of bees dancing ($P > 0.1, 0.7$ and 0.2 for

trials 1 to 3, respectively). This ANOVA found no significant colony effects ($P > 0.7, 0.06$ and 0.8 , respectively), and a significant effect of time of day in trial 1 (more dancing at later hours, $P < 0.001$) but not trial 2 or 3 ($P > 0.2$ and 0.08 , respectively). The ratio of recruits arriving at the feeders during the diffuse-light treatment (Fig. 1b, shaded bars) decreased significantly with distance (tie-corrected Spearman rank correlation coefficient = -0.884 ; $P < 0.02$), showing (like ref. 15) that the effect of dance language vector information is greater further from the colony.

We compared foraging success at natural food sources by measuring the mass of colonies while each foraged under alternating diffuse- and oriented-light treatments. One pair of colonies had diffuse-light and one pair had oriented-light treatments, and we exchanged treatments approximately every 11 days for 8 months (Fig. 1b). The colonies were housed on the agricultural experiment station on the University of California Riverside campus, where bees forage profitably throughout the year, though there is a relative dearth of nectar during autumn. We logged each colony's mass with electronic scales. Colonies often lost mass, when warming the uninsulated glass hive used honey stores faster than foraging replenished them. Figure 2a shows the weight changes for each light-treatment period in the four colonies over the course of the study. We analysed the weight changes with the ANOVA model in Table 1. We found significant effects for season, period nested within season, light treatment and the season $\times$ light treatment interaction. Overall, during periods in which they had oriented dances, these small colonies averaged greater food collection than when their dances were disoriented (mean $\pm$ s.e. mass change = 4 ± 55 g versus -112 ± 45 g per period), but this effect varied significantly in different seasons (Fig. 2b). During summer, colonies gained similar mass under either treatment ($P = 0.22$, Fischer's protected l.s.d.); during autumn, they similarly lost mass regardless of treatment ($P = 0.85$); but during winter, colonies gained mass while under the oriented-light treatment and lost mass while under the diffuse-light treatment ($P = 0.0005$). These effects are visible in the individual periods in Fig. 2a.

These results demonstrate that under natural foraging conditions the communication of distance and direction in the dance language can increase the food collection of honeybee colonies. They also robustly confirm that bees use this directional information in locating the food sources advertised in the dance. However, the study also demonstrates that this advantage does not always hold. In the feeder-recruitment experiments, dance-language vector information increased the number of recruits that found the feeders, but about one-third as many recruits came even without directional information, presumably relying on learning the scent of the feeder from the dancing bees and then searching for that scent in the field^{1,3,4}. During summer and autumn, there was no significant difference in the food collected, as measured by mass changes, when colonies did and did not have directional information from the dance. In a previous study at the same location¹⁸, done during autumn, dances in two adjacent colonies indicated that the colonies were using relatively many different sources. Similar foraging conditions, of relatively equal, dispersed sources, probably occurred during autumn in this study. Under such conditions, foragers

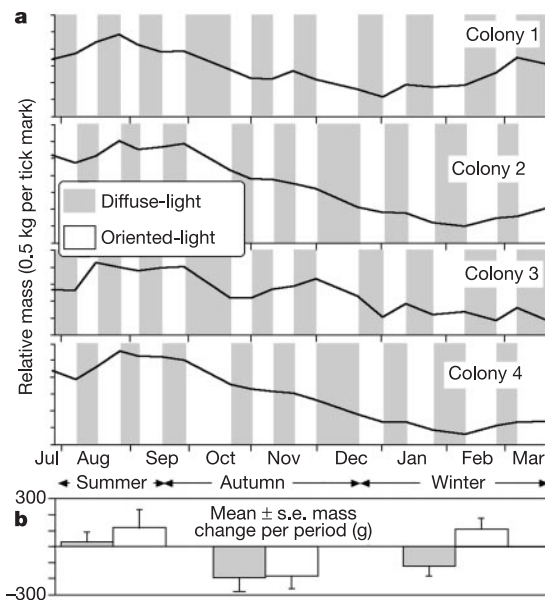


Figure 2 Food collection in free-foraging colonies while receiving diffuse- and oriented-light treatments. **a**, Relative mass of four colonies during each of 19 treatment periods. Shaded regions denote diffuse-light treatment, unshaded denotes oriented-light treatment. **b**, Mass change per treatment period, during periods when colonies had oriented and disoriented dances, for three seasons. Season boundaries are the treatment change nearest the conventional solar dates.

Source	d.f.	SS	MS	F	P
Season	2	0.967	0.484	16.917	0.0001
Period within season	16	4.314	0.270	9.434	0.0001
Colony pair	1	0.098	0.098	3.443	0.0693
Colony within pair	2	0.015	0.008	0.268	0.7663
Light treatment	1	0.264	0.264	9.221	0.0038
Season $\times$ light treatment	2	0.211	0.105	3.688	0.0319
Residual	51	1.458	0.029		

d.f., Degrees of freedom; F, F-statistic; MS, mean square; P, probability; SS, sum of squares.

apparently do just as well at finding a food source without dance vector information, even though they still incur the costs involved in the dance language communication (both from dancing itself and the long time course to decide on and find the advertised food¹¹). However, in conditions like those in winter in this study, communication of food source location does increase a colony's food intake. We note that this difference does not just reflect food abundance. The seasons (summer and autumn) in which oriented dancing was not associated with greater food gain had average mass gains that were greater (summer) and less (autumn) than the season (winter) in which oriented dancing did lead to greater food collection. Temporal variation in the dispersion of natural food sources would be expected to result in such effects. We did not measure this, but just such variation was shown in a different habitat¹⁰.

The evolution of recruitment communication in social insects is presumably steered by whether, for particular colony sizes and habitats, a recruitment mechanism increases food collection more than it decreases it. Different habitats and, as shown here, different seasons within habitats can vary in this payoff. It may be that in the tropical forests of Asia, where the dance language presumably evolved among ancestral *Apis*², food resources from flowering trees were extremely patchy, although this may also vary with time. For honeybees, which seem adapted to a 'boom or bust' economy¹⁹, even relatively short periods within the annual cycle where dancing increases fitness could drive the evolution of dance-language recruitment. Studies of adaptation need to take varying condition-dependent payoffs into account, and examine a range of natural conditions, or run the risk of over-emphasizing or completely missing the impact of particular traits.

Received 27 May; accepted 3 September 2002; doi:10.1038/nature01127.

1. von Frisch, K. *The Dance Language and Orientation of Honeybees* (Harvard Univ. Press, Cambridge, Massachusetts, 1967).
2. Dyer, F. C. The biology of the dance language. *Annu. Rev. Entomol.* **47**, 917–949 (2002).
3. Wenner, A. M. & Wells, P. H. *Anatomy of a Controversy* 57–463 (Columbia Univ. Press, New York, 1990).
4. Gould, J. L. The dance-language controversy. *Q. Rev. Biol.* **51**, 211–244 (1976).
5. Wehner, R. & Rosell, S. in *Experimental Behavioral Ecology And Sociobiology: International Symp., Mainz, West Germany, Oct. 17–19, 1983* (eds Hölldobler, B. & Lindauer, M.) 11–53 (Sinauer, Sunderland, Massachusetts, 1985).
6. Esch, H. E. & Burns, J. E. Distance estimation by foraging honeybees. *J. Exp. Biol.* **199**, 155–162 (1996).
7. Michelsen, A., Andersen, B. B., Storm, J., Kirchner, W. H. & Lindauer, M. How honeybees perceive communication dances, studied by means of a mechanical model. *Behav. Ecol. Sociobiol.* **30**, 143–150 (1992).
8. Kirchner, W. H. Acoustical communication in honeybees. *Apidologie* **24**, 297–307 (1993).
9. Seeley, T. D. *The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies* 69–265 (Harvard Univ. Press, Cambridge, Massachusetts, 1995).
10. Visscher, P. K. & Seeley, T. D. Foraging strategy of honey bee colonies in a temperate deciduous forest. *Ecology* **63**, 1790–1801 (1982).
11. Dornhaus, A. & Chittka, L. Evolutionary origins of bee dances. *Nature* **401**, 38 (1999).
12. Lindauer, M. *Communication Among Social Bees* (Harvard Univ. Press, Cambridge, Massachusetts, 1971).
13. Dyer, F. C. Mechanisms of dance orientation in the Asian honey bee *Apis florea*. *J. Comp. Physiol. A* **157**, 183–198 (1985).
14. Brines, M. L. & Gould, J. L. Bees have rules. *Science* **206**, 571–573 (1979).
15. Kirchner, W. H. & Grasser, A. The significance of odor cues and dance language information for the food search behaviour of honeybees (Hymenoptera: Apidae). *J. Insect Behav.* **11**, 169–178 (1998).
16. Visscher, P. K. *Les Insectes Sociaux. 12th Congress of the International Union for Study of Social Insects* (Université Paris Nord, Paris, 1994).
17. Batschelet, E. *Circular Statistics in Biology* (Academic, New York, 1981).
18. Waddington, K. D., Visscher, P. K., Herbert, T. J. & Richter, M. R. Comparisons of forager distributions from matched honey bee colonies in suburban environments. *Behav. Ecol. Sociobiol.* **35**, 423–429 (1994).
19. Seeley, T. D. & Visscher, P. K. Survival of honeybees in cold climates: the critical timing of colony growth and reproduction. *Ecol. Entomol.* **10**, 81–88 (1985).
20. Sherman, G. *The Role of Dance Language Cues in Recruitment and Food Collection in Honey Bee (*Apis mellifera*) Colonies*. MS thesis, Univ. California Riverside (2002).

Acknowledgements We are grateful to R. Vetter and D. Tanner for assistance at the feeders; R. Beaver for statistical advice; M. Zuk, J. Millar, R. Redak and T. Seeley for feedback; and the Department of Entomology and Academic Senate at UCR for financial support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.K.V. (e-mail: visscher@mail.ucr.edu).

Scotopic colour vision in nocturnal hawkmoths

Almut Kelber, Anna Balkenius & Eric J. Warrant

Department of Cell and Organism Biology, Vision Group, Lund University, Helgonavägen 3, S-22362 Lund, Sweden

Humans are colour-blind at night, and it has been assumed that this is true of all animals. But colour vision is as useful for discriminating objects¹ at night as it is during the day. Here we show, through behavioural experiments, that the nocturnal hawkmoth *Deilephila elpenor* uses colour vision to discriminate coloured stimuli at intensities corresponding to dim starlight (0.0001 cd m⁻²). It can do this even if the illumination colour changes, thereby showing colour constancy—a property of true colour vision systems². In identical conditions humans are completely colour-blind. Our calculations show that the possession of three photoreceptor classes reduces the absolute sensitivity of the eye, which indicates that colour vision has a high ecological relevance in nocturnal moths. In addition, the photoreceptors of a single ommatidium absorb too few photons for reliable discrimination, indicating that spatial and/or temporal summation must occur for colour vision to be possible. Taken together, our results show that colour vision occurs at nocturnal intensities in a biologically relevant context.

Like most vertebrates, humans lose the use of their three spectral classes of cone photoreceptors in dim light, reverting to the colour-insensitive rod system. But many nocturnal insects maintain three spectral classes of photoreceptors at all times of the day^{3,4} and thus have the cellular pre-conditions for scotopic colour vision⁵. *D. elpenor* is a truly nocturnal hawkmoth, feeding on nectar during the darkest hours of the night while hovering in front of flowers⁶. Typical flowers that are pollinated by hawkmoths appear white or bright yellow to the human eye⁷, and it is generally assumed that they attract insects by their high intensity contrast to the background of green leaves. Studies have shown that most white nocturnal flowers absorb ultraviolet light^{8,9}. Because all insect rhodopsins have some sensitivity in the ultraviolet range, these flowers are therefore not maximally bright to an insect eye. Diurnal hawkmoths use trichromatic colour vision for discriminating flowers in bright daylight¹⁰ but, apart from one observation on *Hyles livornica*¹¹, no attempt has been made to test whether nocturnal hawkmoths use their three spectral types of photoreceptor for colour vision at night.

We trained freely flying *D. elpenor* to associate a reward of sugar solution with a colour, either blue or yellow to a human observer. Animals were trained to feed from coloured artificial flowers at a light intensity equivalent to late dusk (0.01 cd m⁻²) and were tested at five different light intensities ranging from mid-dusk (1 cd m⁻²) to dim starlight (0.0001 cd m⁻²). In two tests, different shades of grey were presented together with the training colour (sugar solution absent), in a method similar to that used to show colour vision in honeybees¹². The hawkmoths were able to discriminate blue and yellow from all shades of grey at all light intensities tested, which indicated that they were using colour rather than achromatic cues (Figs 1 and 2, top and middle left).

With only eight shades of grey, however, an eye extremely sensitive to contrast might still be able to use an achromatic signal to make this discrimination. If this were the case, then the moths should distinguish the training blue or yellow from darker and lighter shades of the same colour just as successfully as they distinguished the training colours from grey. We therefore carried out a third test, in which we presented the training colour together

with a lighter and darker shade (produced by adding white or black), and two other colours. The hawkmoths rarely chose the two other colours (Figs 1 and 2, bottom left); however, they did choose the training colour significantly less frequently (G -test, $P < 0.005$) in this test than in both previous tests (Figs 1 and 2), instead choosing the darker or brighter shades. This proved that they were not using achromatic cues to distinguish between different shades of the training colour and therefore could not have used them for distinguishing these colours from grey. Animals trained to yellow even preferred the dark shade of yellow to any shade of grey (three animals, 30/30 correct choices in total, G -test, $P < 0.0001$).

For comparison, six human subjects were asked to make the same colour discriminations under exactly the same conditions. Humans reported that they saw yellow as a colour down to 10^{-3} cd m^{-2} and blue down to 10^{-2} cd m^{-2} but not at lower light intensities. Unlike moths, they were almost always able to discriminate between different shades of a colour (Figs 1 and 2, bottom right; G -tests, $P < 0.0001$; with the exception of yellow under 10^{-4} cd m^{-2} , $P < 0.05$ and blue under 10^{-4} cd m^{-2} , not significant). Humans used achromatic contrast to discriminate blue from bright shades of grey and yellow from dark shades of grey at scotopic light intensities (Figs 1, top right, and 2, middle). But they failed completely when

this cue was not available, for example, when blue was tested against darker shades of grey (Fig. 1, middle right, 10^{-3} and 10^{-4} cd m^{-2}), or when yellow was tested against lighter shades (Fig. 2, top right, left bar).

We conclude that nocturnal hawkmoths use colour vision, rather than achromatic cues, to discriminate between flowers at night, and that their colour vision works at much lower light intensities than that of humans. Additional tests carried out on three striped hawkmoths (*Hyles lineata*) and two bedstraw hawkmoths (*Hyles gallii*) showed that these moths had similar abilities, making it probable that colour vision is common among nocturnal hawkmoths.

Many species of hawkmoth are active at dawn and dusk when the colour of light changes considerably¹³. To find and recognize rewarding flowers independent of the colour of the illuminating light, they require a colour-constant visual system². We tested colour constancy in *D. elpenor* by training them to discriminate between two artificial flowers, one green and one turquoise. Discrimination was tested in two different illumination colours, white and broad-spectrum yellow. In yellow light, the turquoise flower generated almost the same relative quantum catch in the ultraviolet, blue and green photoreceptor classes as the green flower did in white light (Fig. 3a). Without colour constancy moths would confuse these two stimuli.

A group of moths were trained in white light to associate a food reward with the green flower, another group were trained with the turquoise flower. Moths trained to green easily discriminated the green flower from the turquoise one, irrespective of whether they

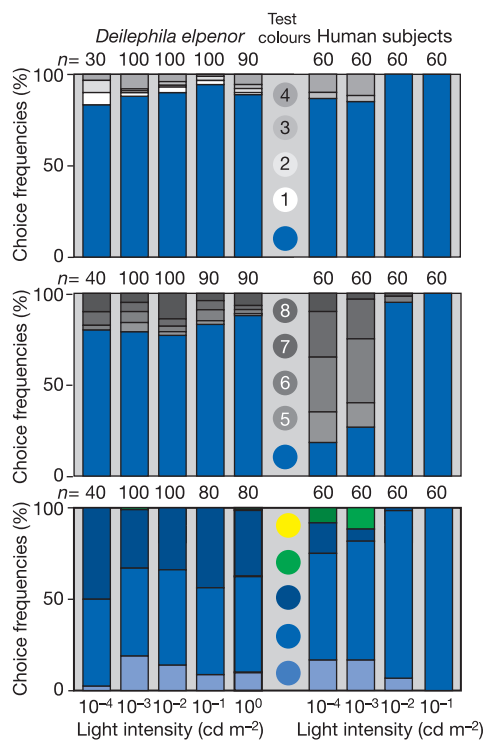


Figure 1 Choice frequencies in tests after training to blue. Each stacked bar diagram gives the choice frequencies to each test colour (indicated by the coloured regions of each bar) at a specific light intensity (see abscissa) for *D. elpenor* (left, number of choices n given above each bar) or six human subjects (right). Three tests with five colours each (presented as disks) were carried out. Moths were able to discriminate blue from all eight shades of grey (numbered) at all intensities (top and middle left, G -test, $P < 0.0001$) but not from lighter and darker shades of blue (bottom left, G -test, $P > 0.1$). Humans discriminated blue from bright shades of grey at all light intensities (top right, G -test, $P < 0.0001$) and from dark shades of grey at mesopic light intensities (middle right, G -test, $P < 0.0001$). They could not discriminate blue from dark shades of grey at scotopic intensities (10^{-3} and 10^{-4} cd m^{-2}): correct choice frequencies were significantly lower than 50% (G -test, $P < 0.0001$) and did not differ from 20% (G -tests, $P > 0.2$) under these conditions. Discrimination between the three shades of blue was significant at 10^{-1} cd m^{-2} , 10^{-2} cd m^{-2} (G -test, $P < 0.0001$) and 10^{-3} cd m^{-2} ($P < 0.05$) but not at 10^{-4} cd m^{-2} ($P < 0.1$; bottom right).

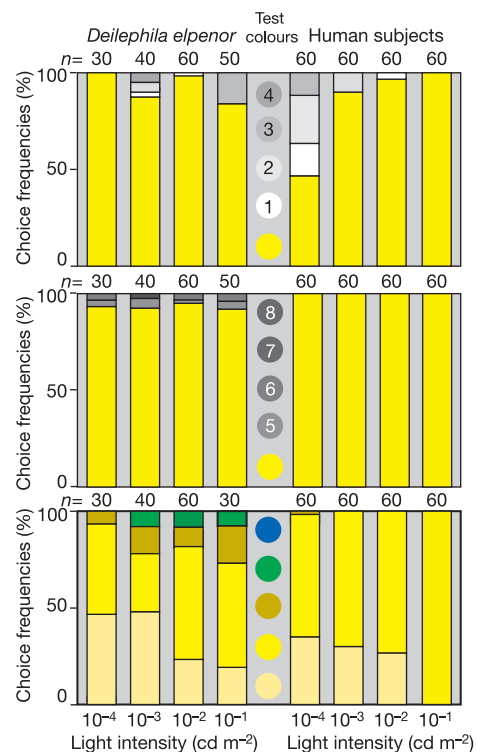


Figure 2 Choice frequencies in tests after training to yellow. See Fig. 1 for details. *D. elpenor* chose yellow more frequently than all eight shades of grey (numbered) at all light intensities (G -test, $P < 0.0001$) but did not discriminate between different shades of yellow ($P > 0.1$). Humans discriminated yellow from dark shades of grey (G -test, $P < 0.0001$) and from the brighter and darker shades of yellow at all light intensities (G -test, $P < 0.0001$ for 10^{-1} cd m^{-2} and 10^{-2} cd m^{-2} , $P < 0.005$ for 10^{-3} cd m^{-2} , and $P < 0.05$ for 10^{-4} cd m^{-2}) but could not discriminate yellow from lighter shades of grey at 10^{-4} cd m^{-2} (G -test, $P > 0.2$).

were illuminated by white (Fig. 3b, left) or yellow (Fig. 3b, right) light. The same result was found for moths trained to turquoise flowers (Fig. 3c): turquoise flowers were chosen almost exclusively, regardless of the illumination spectrum. Hawkmoths chose the rewarding colour with high precision in both illuminations (G -tests, $P < 0.0001$), proving that they have colour constancy. The result also shows high flexibility in learning: *D. elpenor* could learn to associate a reward with a colour that is not typical of flowers, such as green.

Using optical, anatomical and physiological data, we calculated the number of photons absorbed by the nine photoreceptors in a single ommatidium when moths viewed artificial flowers and their backgrounds at the dimmest light intensity used in the experiments (0.0001 cd m^{-2}). *D. elpenor* has two different types of ommatidium in the eye: both types have seven proximal green receptors, but one has two distal blue receptors, and the other has two distal ultraviolet receptors. The photon catches for the three shades of blue and three shades of grey are shown in Fig. 4. During one visual integration time, quantum catches for all six stimuli were between 1 and 16 photons per photoreceptor type (Fig. 4). In this range, photon shot noise ($\sqrt{N}$, where N is the number of absorbed photons¹⁴) is the dominant source of noise^{15,16}, resulting in a signal-to-noise ratio of between 1 and 4. Reliable discrimination of blue from grey is impossible under these conditions¹⁴, and this is especially true for the training blue and the medium shade of grey (shade 6) shown in Fig. 4. For these two stimuli, the quantum catches were almost identical. Therefore, we have to assume that moths use spatial and/or temporal summation^{17,18} to increase the signal-to-noise ratio; otherwise, the colour discrimination in dim starlight that we have shown would not be possible. Our calculations also show that to see colour at night, hawkmoths sacrifice absolute sensitivity, a strong indication that colour vision is highly relevant to their ecology (data not shown).

As the quantum catches for the ultraviolet photoreceptor were especially low (Fig. 4), we tested whether these photoreceptors contribute to colour vision at all. We trained three *D. elpenor* moths to discriminate a white stimulus that reflected ultraviolet light from one that absorbed it. The ultraviolet-absorbing white stimulus was selected as the reward colour for training, and in tests (without reward at 0.01 cd m^{-2}) the moths chose the reward colour exclusively (23/23 correct choices for the reward colour, G -test, $P < 0.0001$). This indicates that *D. elpenor* may possess trichromatic colour vision, although further experiments would be required to establish true trichromacy.

To our knowledge, this is the first report of an animal using

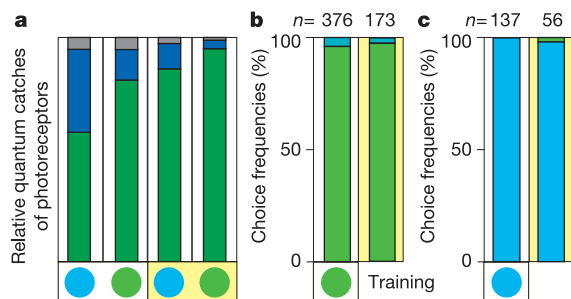


Figure 3 Colour constancy in *D. elpenor*. **a**, Relative receptor quantum catches calculated by equation (1) in the three photoreceptor channels generated by turquoise and green stimuli illuminated by white (two left columns) or yellow (two right columns) light. In each stacked bar, grey regions represent ultraviolet receptors, blue regions represent blue receptors, and green regions represent green receptors. **b**, **c**, Moths trained to green stimuli (**b**) and turquoise stimuli (**c**) chose correctly in tests under two different illuminations, 'white' (left) and 'yellow' (right). Other conventions are the same as in Fig. 1.

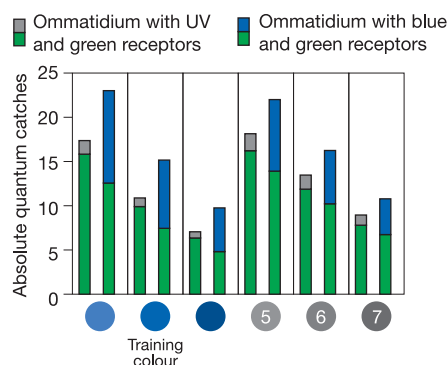


Figure 4 Quantum catches of all *D. elpenor* photoreceptors in the two types of ommatidium (two ultraviolet and seven green receptors, two blue and seven green receptors) viewing six of the test colours (including the blue training colour) at the dimmest light intensity used ($10^{-4} \text{ cd m}^{-2}$). Colour conventions in each stacked bar are the same as in Fig. 3a. Lights were measured using a calibrated S2000 Ocean Optics spectrometer and an International Light IL1700 photometer in units of photons $\text{s}^{-1} \text{ nm}^{-1} \text{ m}^{-2} \text{ steradian}^{-1}$, and quantum catches were calculated using equation (1).

colour vision for object discrimination at nocturnal intensities. Nocturnal hawkmoths have large and sensitive superposition compound eyes. These, and presumably neural summation mechanisms, allow them to maintain colour vision even at light intensities in which humans and insects with apposition eyes, such as bees¹⁹, become colour-blind. The only earlier indication of colour discrimination in dim light comes from goldfishes²⁰, in which heart rate conditioning was used to suggest that an opponent interaction of rods and red cones may enable fishes to discriminate red and green lights. But the biological relevance of this mechanism, if any, is not known. Nocturnal hawkmoths such as *D. elpenor* obviously rely on colour discriminations for choosing flowers, just as diurnal species do¹⁰. □

Methods

Animals and behavioural tests

We obtained *D. elpenor*, *H. lineata* and *H. gallii* from several breeders. All larvae were fed on their natural food plants to avoid visual pigment deficiencies. Animals were kept under a 12 h/12 h dark/light regime for the duration of the experiment, and training started 2 d after eclosion, during the dark phase. The experimental cage was 50 cm long, 60 cm high, 70 cm wide and illuminated from above by a broad-spectrum ultraviolet-visible high pressure mercury lamp (Leitz) that could be dimmed by quartz neutral density filters (and a broad-spectrum yellow Schott FG-13 filter). Coloured paper stimuli, generated using an Epson colour printer and Epson inkjet paper, were circular, 30 mm in diameter, and were presented vertically on the cage wall. We used regular white printer paper for ultraviolet-absorbing white and laboratory filter paper as ultraviolet-reflecting white. A light grey background was used in most experiments, but training to white stimuli required a black background otherwise the animals would not approach.

We provided a 20% sugar solution as a reward in the reversed tip of a syringe that could be accessed through a 3-mm-diameter hole in the centre of the training stimulus. One animal was released at a time. During training, it approached the stimulus, extended its long (2.5 cm) proboscis, and probed the stimulus until it found the hole leading to the sugar reservoir. During tests, no reward was present and the moth chose between different colour disks presented simultaneously, by touching them with the proboscis. The first colour disk probed during an approach was taken as the moth's choice. The positions of colours were changed regularly in a pseudorandom manner to avoid any influence from spatial cues. Ten animals were trained to associate blue with a reward, and six were trained to yellow. Three animals were trained to discriminate ultraviolet-absorbing from ultraviolet-reflecting white, and in the constancy test ten animals were trained to green and turquoise, respectively. Six adult humans with normal colour vision were asked to make the same discriminations as the moths. Ten choices by each subject were recorded at each light intensity.

We used G -tests to decide whether the choice frequency for the training colour was significantly higher than 50%. If it was, we concluded that the moths or human subjects were able to discriminate the training colour from each of the other test colours.

Light measurements and quantum catches calculation

Light reflected from stimuli was measured using an International Light IL 1700 radiometer

and a calibrated Ocean Optics S2000 spectrometer. We assumed that the eyes of *D. elenor* have two ommatidial types: one containing two distal ultraviolet receptors and seven proximal green receptors; and one with two distal blue receptors and seven proximal green receptors. Evidence for this assumption comes from anatomical data on *D. elenor*²¹ and on the closely related species *Manduca sexta* (R. White, personal communication). We calculated the number of photons, N , that are absorbed by each type of photoreceptor in one ommatidium per integration time of the photoreceptor, using optical and electrophysiological data for *D. elenor* (E.W., unpublished data) and the following equation^{18,22}:

$$N = 1.13 \left(\frac{\pi}{4} \right) n \Delta \rho^2 D^2 \kappa \tau \Delta t (1 - e^{-k R_i(\lambda) l}) I(\lambda) \quad (1)$$

where $I(\lambda)$ is the stimulus light intensity in photons $\text{s}^{-1} \text{nm}^{-1} \text{m}^{-2} \text{steradian}^{-1}$, n is the number of effective facets in the superposition aperture²³, 568; $\Delta \rho$ is the photoreceptor acceptance angle, 3.0° ; D is the diameter of a facet lens, $29 \mu\text{m}$; κ is the quantum efficiency of transduction, 0.5 ; τ is the transmission of the eye media, 0.8 ; Δt is the integration time of a photoreceptor, 0.036 s ; k is the absorption coefficient of the rhabdom, $0.0067 \mu\text{m}^{-1}$, l is the rhabdom length, doubled by tapetal reflection, $414 \mu\text{m}$; and $R_i(\lambda)$ is the spectral sensitivity of photoreceptor i (where $i = 1, 2, 3$), calculated from the recorded sensitivity maxima^{3,4} using the Stavenga–Smits–Hoenders rhodopsin template²⁴. Relative quantum catches in each of the three receptor channels (Fig. 3) were obtained by dividing their absolute quantum catches by the sum of quantum catches in all three channels. For this purpose, the average of the quantum catches of the green receptors in the two types of ommatidium of *D. elenor* was used.

Received 23 March; accepted 5 August 2002; doi:10.1038/nature01065.

- Kelber, A., Vorobyev, M. & Osorio, D. Animal colour vision—behavioural tests and physiological concepts. *Biol. Rev.* (in the press).
- Neumeyer, C. *Perceptual Constancy: Why Things Look as They Do* (eds Walsh, V. & Kulikowski, J.) 323–351 (Cambridge Univ. Press, Cambridge, 1998).
- Höglund, G., Hamdorf, K. & Rosner, G. Trichromatic visual system in an insect and its sensitivity control by blue light. *J. Comp. Physiol.* **86**, 265–279 (1973).
- Schwemer, J. & Paulsen, R. Three visual pigments in *Deilephila elenor* (Lepidoptera, Sphingidae). *J. Comp. Physiol.* **86**, 215–229 (1973).
- Schlecht, P. Colour discrimination in dim light. An analysis of the photoreceptor arrangement in the moth *Deilephila*. *J. Comp. Physiol.* **129**, 257–267 (1979).
- Pittaway, A. R. *The Hawkmoths of the Western Palaearctic* (Harley Books, Colchester, 1993).
- Vogel, S. *Blütenbiologische Typen als Elemente der Sipplengliederung* (G. Fischer, Jena, 1954).
- White, R. H., Stevenson, R. D., Bennett, R. R., Cutler, D. E. & Haber, W. A. Wavelength discrimination and the role of ultraviolet vision in the feeding behaviour of hawkmoths. *Biotropica* **26**, 427–435 (1994).
- Raguso, R. A. & Willis, M. A. *Proceedings of the Third International Symposium on Butterfly Ecology and Evolution* (eds Watt, W. B., Boggs, C. & Ehrlich, P. R.) (Univ. of Chicago Press, Chicago, in the press).
- Kelber, A. & Hénique, U. Trichromatic colour vision in the hummingbird hawkmoth, *Macroglossus stellatarum*. *J. Comp. Physiol.* **A 184**, 535–541 (1999).
- Knoll, F. Lichtsinn und Blütenbesuch des Falters von *Deilephila livornica*. *Z. vergl. Physiol.* **2**, 328–380 (1926).
- Frisch, K. von Der Farbensinn und Formensinn der Biene. *Zool. Jb. Abt. Zool. Physiol.* **35**, 1–188 (1914).
- Minnaert, M. G. J. *Light and Color in the Outdoors* (Springer, New York, 1993).
- Land, M. F. *Handbook of Sensory Physiology* (ed. Autrum, H.) **Vol. VII/6B** 471–592 (Springer, Berlin, 1981).
- Rose, A. The relative sensitivities of television pickup tubes, photographic film and the human eye. *Proc. Inst. Radio Eng.* **NY 30**, 293–300 (1942).
- De Vries, H. The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica* **10**, 553–564 (1943).
- Warrant, E., Porombka, T. & Kirchner, W. H. Neural image enhancement allows honeybees to see at night. *Proc. R. Soc. Lond. B* **263**, 1521–1526 (1996).
- Warrant, E. J. Seeing better at night: life-style, eye design and the optimum strategy of spatial and temporal summation. *Vision Res.* **39**, 1611–1630 (1999).
- Menzel, R. Achromatic vision in the honeybee at low light intensities. *J. Comp. Physiol.* **A 141**, 389–393 (1981).
- Powers, M. K. & Easter, S. S. Wavelength discrimination by the goldfish near absolute visual threshold. *Vision Res.* **18**, 1149–1154 (1978).
- Welsch, B. Ultrastruktur und funktionelle Morphologie der Augen des Nachtfalters *Deilephila elenor* (Lepidoptera, Sphingidae). *Cytobiologie* **14**, 378–400 (1977).
- Warrant, E. J. & Nilsson, D.-E. Absorption of white light in photoreceptors. *Vision Res.* **38**, 195–207 (1998).
- Warrant, E. J. & McIntyre, P. D. Screening pigment, aperture and sensitivity in the dung beetle superposition eye. *J. Comp. Physiol.* **A 167**, 804–816 (1990).
- Stavenga, D. G., Smits, R. P. & Hoenders, B. J. Simple exponential functions describing the absorbance bands of visual pigment spectra. *Vision Res.* **33**, 1011–1017 (1993).

Acknowledgements We thank D.-E. Nilsson for critically reading the manuscript; M. Pfaff for insights into hawkmoth behaviour; the human volunteers for participating in experiments; and everybody in the Vision Group for inspiration. We are grateful for the financial support of the Swedish Research Council.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.K. (e-mail: almut.kelber@zool.lu.se).

Antagonistic pathways in neurons exposed to body fluid regulate social feeding in *Caenorhabditis elegans*

Juliet C. Coates*† & Mario de Bono*

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Wild isolates of *Caenorhabditis elegans* can feed either alone or in groups^{1,2}. This natural variation in behaviour is associated with a single residue difference in NPR-1, a predicted G-protein-coupled neuropeptide receptor related to Neuropeptide Y receptors². Here we show that the NPR-1 isoform associated with solitary feeding acts in neurons exposed to the body fluid to inhibit social feeding. Furthermore, suppressing the activity of these neurons, called AQR, PQR and URX, using an activated K^+ channel, inhibits social feeding. NPR-1 activity in AQR, PQR and URX neurons seems to suppress social feeding by antagonizing signalling through a cyclic GMP-gated ion channel encoded by *tax-2* and *tax-4*. We show that mutations in *tax-2* or *tax-4* disrupt social feeding, and that *tax-4* is required in several neurons for social feeding, including one or more of AQR, PQR and URX. The AQR, PQR and URX neurons are unusual in *C. elegans* because they are directly exposed to the pseudocoelomic body fluid³. Our data suggest a model in which these neurons integrate antagonistic signals to control the choice between social and solitary feeding behaviour.

In addition to aggregating together, *C. elegans* social feeders accumulate at the border of a bacterial food lawn (termed bordering) and do not strongly downregulate locomotion in response to food. In contrast, well-fed solitary feeders show no preference for the lawn border, and slow down on encountering food². All these behavioural differences are associated with natural variation at residue 215 of NPR-1: social strains possess phenylalanine at this position (NPR-1 215F), whereas solitary strains possess valine (NPR-1 215V). Animals lacking *npr-1* activity are strongly social, indicating that *npr-1* functions to repress social feeding. Both natural isoforms of NPR-1 can repress social feeding, but *npr-1* 215V does so more potently².

To identify cells that express NPR-1, we made a transgene expressing NPR-1 215V tagged with green fluorescent protein (GFP) from the *npr-1* promoter (Fig. 1a; see also Methods). This transgene completely suppressed social feeding in *npr-1(ad609)* mutant animals. *npr-1(ad609)* is a strong loss-of-function allele of *npr-1* (Fig. 1b, c). Rescued transgenic animals showed NPR-1::GFP fluorescence predominantly in the nervous system (Fig. 1d–g). We identified the NPR-1::GFP-expressing cells in the head as the sensory neurons AQR, ASE, ASG, ASH (L4/adult stages only), URX, IL2L/R and OLQ (with its socket and sheath cells), the interneurons AUA and SAAD, the motor neurons RMG and SMBD, and the pharyngeal neuron M3. In the ventral nerve cord, expression was observed in the GABA (γ -aminobutyric acid)-containing VD and DD motoneurons (also reported in ref. 4). In the tail, expression was seen in the sensory neurons PQR, PHA and PHB. We also tentatively identified the interneurons RIV, RIG and SDQ as expressing NPR-1::GFP. For most neurons, expression was seen throughout post-embryonic life, usually in both partners of bilaterally symmetrical neuron pairs. NPR-1::GFP fluorescence was visible on neuron cell bodies, axons and dendrites (Fig. 1d–g), consistent with NPR-1 acting as a neuropeptide transmitter receptor. In contrast, *C. elegans* chemosensory receptors are localized to sensory cilia^{5,6}. NPR-1::GFP fluorescence was also seen in a muscle

† Present address: Department of Plant Sciences, University of Cambridge, Cambridge, UK.

in the terminal bulb of the pharynx, and sometimes in the excretory duct cell and excretory canal.

NPR-1 215V may suppress social feeding by regulating neural development. To examine whether NPR-1 activity in adults is sufficient to induce solitary feeding we expressed NPR-1 215V in *npr-1* animals under the control of the *hsp-16* heat-inducible promoter. This promoter expresses predominantly in neurons and shows no expression in the absence of heat shock⁷. Social adult *npr-1(ad609)* animals became solitary feeders for at least 4 h within 60 min of heat-shock-induced *npr-1* expression (Fig. 2a), and showed reduced bordering behaviour (data not shown). This suggests that *npr-1* acts acutely to suppress social feeding, although we cannot rule out the possibility that expression from the heat shock promoter disrupted aggregation and bordering nonspecifically.

To identify the neurons where NPR-1 215V functions to inhibit social feeding, we expressed the receptor in small subsets of NPR-1-expressing neurons using heterologous promoters in a social *npr-1(ad609)* mutant background. Expression of NPR-1 215V specifically in the AQR, PQR and URX sensory neurons, using the *gcy-32* promoter⁸, suppressed both aggregation and bordering of *npr-1(ad609)* mutant animals (Fig. 2b, c). Stronger suppression of social feeding was obtained if *npr-1* was expressed in the AUA interneuron in addition to AQR, PQR and URX, using the *flp-8* promoter (Fig. 2b, c; C. Li, personal communication; see also Methods). Weak aggregation and bordering remained even with expression in the AQR, PQR, URX and AUA neurons, suggesting that *npr-1* may also function in additional cells to inhibit social feeding. AQR, PQR and URX are the only *C. elegans* neurons exposed directly to the body fluid of the animal apart from the CEP neurons³. AUA is a principal synaptic target of URX³.

Our results implicated the body cavity neurons AQR, PQR and URX in the regulation of social feeding. These neurons express the *eag*-type potassium channel EGL-2 (ref. 9). Gain-of-function (gf) alleles of this gene, *egl-2(n693)* and *egl-2(n2656)*, encode K⁺ channels that open at inappropriately negative voltages and inhibit the activity of the neurons in which they are expressed⁹. We tested whether *egl-2(gf)* mutations could suppress social feeding behaviour. Both *egl-2(n693)* and *egl-2(n2656)* mutations suppressed strongly the aggregation and bordering behaviours of *npr-1(ad609)* animals (Fig. 3a, b, and data not shown), but did not suppress the high locomotory activity of *npr-1(ad609)* animals on food, as *egl-2; npr-1* double mutant animals move faster than *egl-2* single mutants (Fig. 3c). *egl-2* is expressed in seven other neurons apart from AQR, PQR and URX⁹. To test whether EGL-2(gf) activity in AQR, PQR and URX alone is sufficient to suppress social feeding, we expressed *egl-2(gf)* complementary DNA specifically in these neurons in *npr-1(ad609)* animals using the *gcy-32* promoter. We monitored *egl-2* expression by placing sequences encoding GFP downstream of the *egl-2* cDNA in an operon (see Methods). Expression of EGL-2(gf) in AQR, PQR and URX neurons recapitulated the effects of *egl-2(gf)* alleles on *npr-1* animals. It was sufficient to suppress strongly the aggregation and bordering phenotypes of *npr-1(ad609)* animals, but did not suppress *npr-1(ad609)* high locomotory activity on food (Fig. 3a–c). *npr-1* mutant animals expressing EGL-2(gf) from the *gcy-32* promoter showed no overt differences to wild-type solitary N2 animals, and displayed wild-type chemotaxis to odours and avoidance of high osmotic tensions (data not shown).

The body cavity neurons AQR, PQR and URX may be required to generate stimuli that allow social animals to aggregate together.

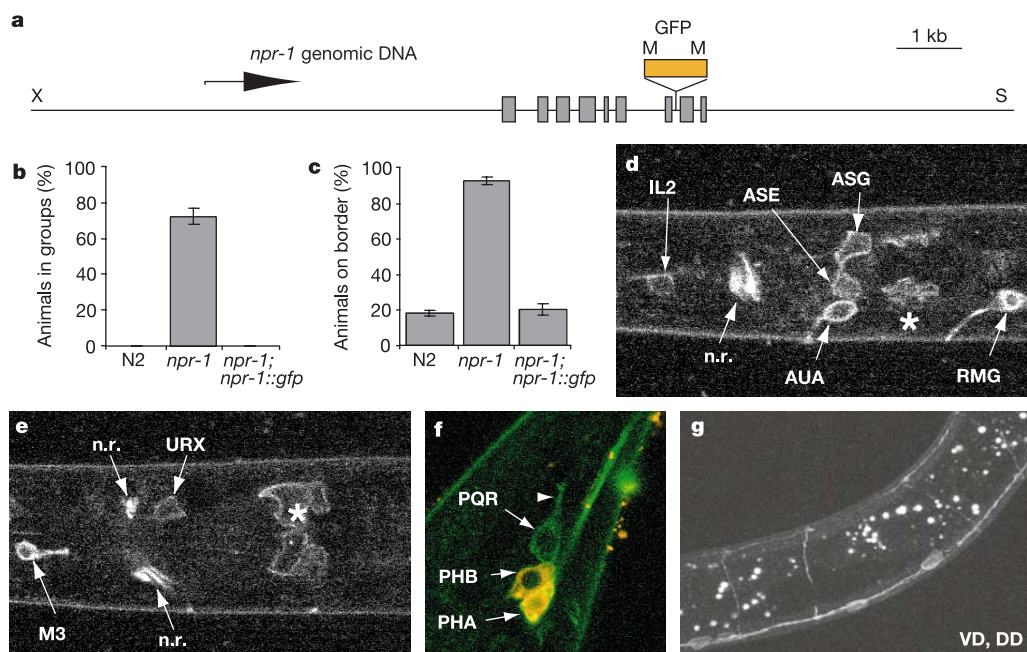


Figure 1 A functional *npr-1::gfp* fusion gene is expressed in approximately 20 neuron types. **a**, The *npr-1::gfp* construct. X, *Xho*I; S, *Sad*I; M, *Mlu*I. *npr-1::gfp* restores N2-like solitary feeding to *npr-1(ad609)* animals. **b**, **c**, For both aggregation (**b**) and bordering (**c**) behaviours, $P < 0.001$ for a comparison between *npr-1* and *npr-1; npr-1::gfp*, and $P > 0.5$ for a comparison between N2 and *npr-1; npr-1::gfp*. Shown are means $\pm$ s.e.m. ($n = 5$ assays). **d–g** *npr-1::gfp* expression. **d**, Left lateral section of the head showing expression in the sensory neurons ASE, ASG and IL2, the interneuron AUA, the motor neuron RMG, and pharyngeal muscle (asterisk; also seen in **e**). n.r., nerve ring.

e, Medial section of the head showing expression in the body cavity neuron URX and the pharyngeal neuron M3. **f**, Section from the left side of the tail showing expression in the phasmid sensory neurons PHA and PHB, and the body cavity neuron PQR (green). Phasmid cell identity was confirmed by co-staining with Dil (red). The arrowhead indicates the sensory cilium of PQR projecting into the body cavity. **g**, Merged sections from the mid-body showing expression in the VD and DD motor neurons in the ventral cord. Anterior is left and dorsal is up in panels **d** and **e**; anterior is bottom left and dorsal is top left in panels **f** and **g**.

Alternatively, these neurons may regulate the response to aggregation stimuli. *npr-1* animals in which AQR, PQR and URX activity was inhibited by expressing the *gcy-32::egl-2(gf)* transgene remained solitary in the presence of marked social *npr-1* animals (Fig. 3d). This result suggests that these neurons regulate responses to aggregation-promoting signals, although an extra involvement in the production of aggregation stimuli cannot be excluded.

Cyclic nucleotide-gated ion channels are important transducers of sensory signals in both vertebrates and invertebrates^{10,11}. In *C. elegans* the *tax-2* and *tax-4* genes encode β - and α -subunits, respectively, of a cyclic GMP-gated cation channel that is required for taste, thermosensation, and some forms of olfaction^{12–14}. *tax-2* and *tax-4* are expressed in overlapping sets of about 12 sensory neurons, including AQR, PQR and URX for *tax-4* (ref. 14 and data not shown), and AQR and PQR for *tax-2* (ref. 12, and C. Bargmann, personal communication). We found that loss-of-function mutations in either *tax-2* or *tax-4* suppressed aggregation and bordering behaviours of *npr-1(ad609)* animals, and restored a slowing response to food (Fig. 4a–c). A promoter mutation that disrupts *tax-2* expression in the AQR, ASE, AFD and BAG sensory neurons¹², *tax-2(p694)*, also disrupted social feeding (Fig. 4a–c). This suggests that *tax-2* is required in one or more of these four neurons to promote social feeding behaviour. Interestingly, cGMP

has been implicated in regulating foraging behaviour in *Drosophila* and bees^{15,16}.

To investigate the neurons in which TAX-4 acts to promote social feeding, a *tax-4* cDNA¹⁴ was introduced into *tax-4(p678); npr-1(ad609)* solitary animals under the control of neuron-specific promoters. We monitored *tax-4* expression using GFP in an artificial operon, as in the *egl-2* expression experiments. Expression of *tax-4* from the *tax-4* promoter¹⁴ restored aggregation and bordering behaviours to *tax-4; npr-1* animals (Fig. 4d, e). Full rescue was also obtained by combined expression of *tax-4* from the *gcy-32* promoter and the *odr-4* promoter (Fig. 4d, e); however, *tax-4* expression from the *gcy-32* promoter alone, or from the *odr-4* promoter alone, did not restore aggregation or strong bordering behaviours to *tax-4; npr-1* animals (Fig. 4d, e, and data not shown). The *odr-4* promoter drives expression in 12 neurons⁶, 6 of which also express *tax-4* (ref. 14). These six neurons are ASG, ASI, ASJ, ASK, AWB and AWC. Our results suggest that social feeding requires *tax-4* activity in one or more of the AQR, PQR and URX body cavity

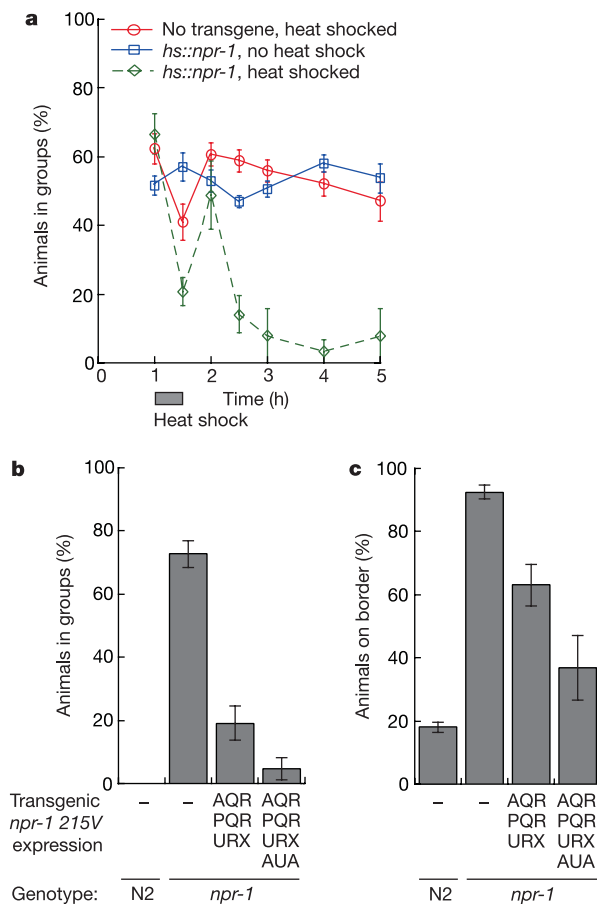


Figure 2 Expression of *npr-1::gfp* from a heat-shock-inducible promoter in adults, or specifically in the AQR, PQR and URX neurons, suppresses social feeding in *npr-1* animals. **a**, Social *npr-1(ad609)* animals expressing *npr-1::gfp* from the *hsp-16* heat-shock-inducible promoter become solitary feeders after heat shock. **b**, **c**, Expression of *npr-1::gfp* specifically in the AQR, PQR and URX body cavity neurons using the *gcy-32* promoter suppresses aggregation (**b**) ($P < 0.001$) and bordering (**c**) ($P < 0.01$) behaviours of *npr-1(ad609)* animals. Additional expression in the AUA neurons using the *flp-8* promoter enhances suppression. Shown are means $\pm$ s.e.m. ($n \geq 4$ assays).

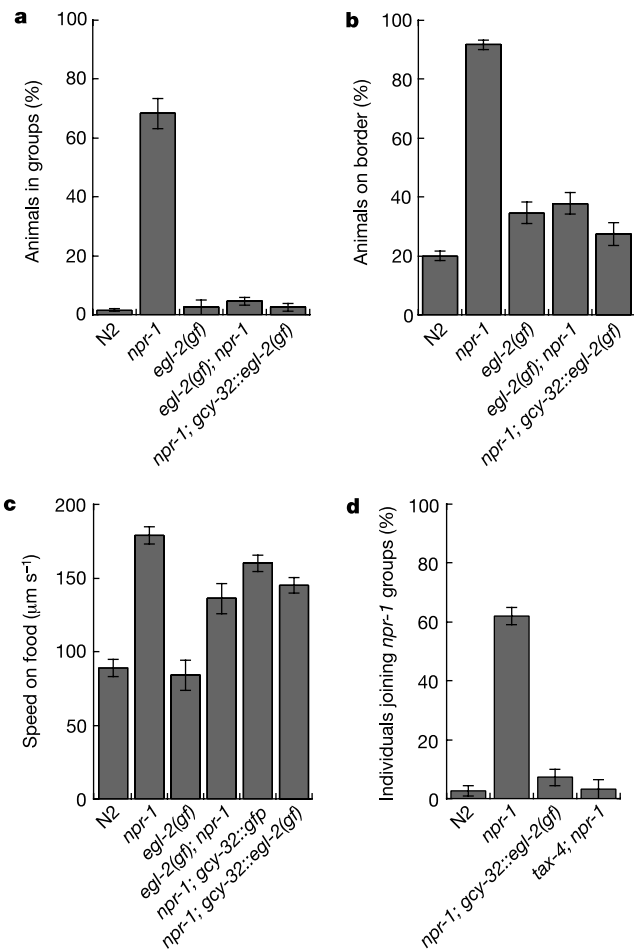


Figure 3 Targeted expression of the EGL-2(gf) dominantly active K^+ channel in the AQR, PQR and URX body cavity neurons suppresses social feeding. **a**, **b**, The *egl-2(n693gf)* allele suppresses aggregation (**a**) and bordering (**b**) of *npr-1(ad609)* animals ($P < 0.001$). Expression of *egl-2(gf)* cDNA⁹ specifically in the AQR, PQR and URX neurons, using the *gcy-32* promoter recapitulates this suppression ($P < 0.001$). **c**, The *gcy-32::egl-2(gf)* transgene does not suppress the high locomotory activity of *npr-1(ad609)* animals on food ($P > 0.05$ for a comparison between *npr-1; gcy-32::gfp* and *npr-1; gcy-32::egl-2(gf)*). **d**, *npr-1(ad609)* animals expressing the *gcy-32::egl-2(gf)* transgene do not join groups of marked *npr-1(ad609)* animals. *tax-4(p678); npr-1(ad609)* solitary animals also fail to join *npr-1* groups. $P < 0.001$ for comparisons between *npr-1* and *npr-1; gcy-32::egl-2(gf)* or *tax-4; npr-1* animals. Shown are means $\pm$ s.e.m. **a**, **b**, $n \geq 4$ assays; **c**, $n \geq 54$ animals; **d**, $n \geq 46$ animals.

neurons, but also in one or more *tax-4*-expressing neurons that also express *odr-4*. Furthermore, *tax-4* activity is required at least in part to mediate a response to signals that promote aggregation, as *tax-4*; *npr-1* animals do not join groups of *npr-1* animals (Fig. 3d).

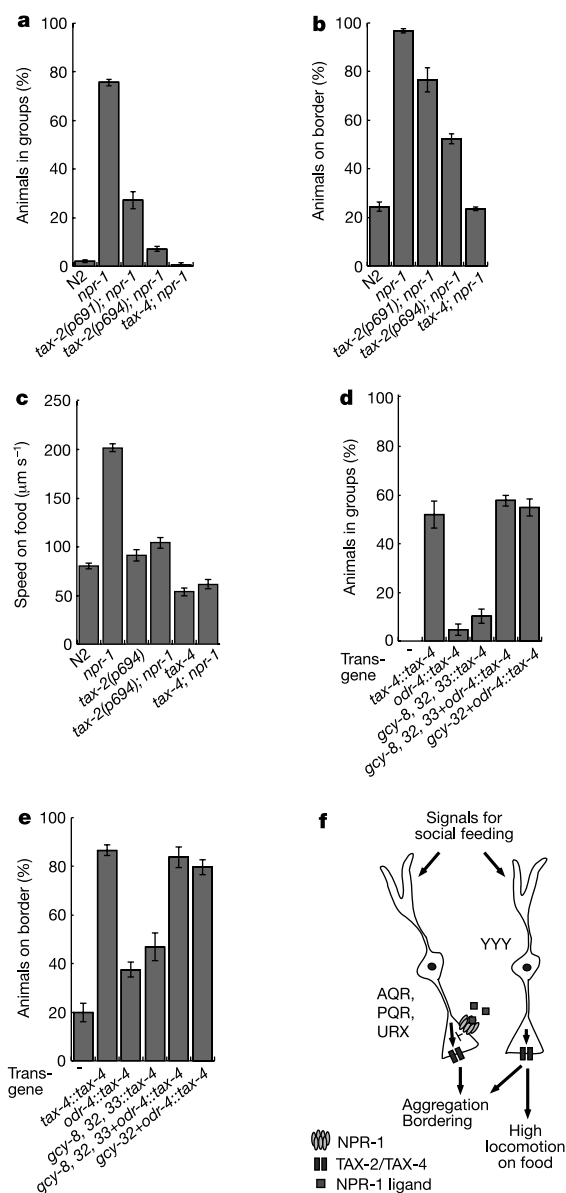


Figure 4 The cGMP-gated ion channel subunits *tax-2* and *tax-4* mediate social feeding. **a–c**, *tax-2* and *tax-4* disrupt aggregation (**a**), bordering (**b**) and high locomotory activity (**c**) of *npr-1(ad609)* animals on food. *tax-4(p678)*, and *tax-2(p691)* are predicted to be strong loss-of-function alleles; *tax-2(p694)* is a promoter mutation that disrupts *tax-2* expression in the AQR, AFD, ASE and BAG neurons (ref. 12 and C. Bargmann, personal communication). **d**, **e**, Combined expression of *tax-4* cDNA from the *gcy-32* and *odr-4* promoters, but not from either promoter alone, restores aggregation (**d**) and bordering (**e**) behaviour of *tax-4(p678)*; *npr-1(ad609)* animals to the level obtained by expressing *tax-4* cDNA from the *tax-4* promoter. The *gcy-8* and *gcy-33* promoters direct expression in the AFD and BAG sensory neurons, respectively⁸, and were used to show that *tax-4* expression in these neurons¹⁴ is not required for social feeding. In both **d**, and **e**, $P < 0.001$ for a comparison between *tax-4::tax-4* and either *gcy-8, -32, -33::tax-4* or *odr-4::tax-4*. $P > 0.4$ for a comparison between *tax-4::tax-4* and *gcy-32::tax-4* plus *odr-4::tax-4*. Shown are means $\pm$ s.e.m. ($n \geq 4$ assays). **f**, One model for the regulation of social feeding by *npr-1* and *tax-2/tax-4*. Activation of the TAX-2/TAX-4 cGMP-gated ion channel in the AQR, PQR and URX body cavity neurons and also in the unidentified YYY neuron(s) activates social feeding behaviour. This activation can be inhibited by NPR-1 activity in AQR, PQR and URX neurons.

Our data suggest a model in which antagonistic signalling pathways in one or more of the body cavity neurons AQR, PQR and URX regulate the choice between social and solitary feeding behaviour (Fig. 4f). One pathway, mediated by *npr-1*, inhibits aggregation and bordering behaviours. A second pathway, which requires the cGMP-gated ion channel subunit TAX-4, promotes social feeding when NPR-1 activity is low. TAX-4 is also required in an additional sensory neuron(s) for social feeding. Designated YYY in our model, this *tax-4*-expressing neuron is probably one or more of the ASG, ASI, ASJ, ASK, AWB and AWC chemosensory neurons. The inhibitory action of NPR-1 may be similar to that of mammalian NPY receptors, which act presynaptically to inhibit the activity of both hippocampal neurons¹⁷ and hypothalamic proopiomelanocortin-containing neurons that regulate appetite¹⁸. Although we have not demonstrated that AQR, PQR and URX all regulate social feeding, these neurons share the unusual feature of contacting the pseudocoelomic fluid, or body fluid, of the animal directly³. In nematodes, the body fluid is a complex solution that seems to function analogously to blood¹⁹. One possibility is that TAX-4 and NPR-1 regulate social feeding behaviour in *C. elegans* by transducing antagonistic signals from the body fluid. The cognate ligand for NPR-1 is unknown, but a range of candidate neuropeptide ligands has been described in the *C. elegans* genome²⁰. Humoral signals are known to regulate mammalian feeding²¹. They can also regulate social behaviours; for example the level of juvenile hormone in bees correlates with division of labour in the hive²². Social foraging theory suggests that feeding in groups involves a trade-off between benefits and penalties²³. Antagonistic inputs regulating aggregation may enable animals to match their social strategy to their environment, both over the course of an individual's life and over an evolutionary timescale. □

Methods

Strains

Nematodes were grown under standard conditions²⁴. Double-mutant strains were constructed using balancer chromosomes in an N2 background. Strains used or generated in this study are listed in Supplementary Information. Transgenic strains were made using *lin-15* (pJMZ, 30 ng μl^{-1})²⁵ as a co-injection marker following standard methods²⁶. Animals injected were either *npr-1(ad609) lin-15(n765ts)* or *tax-4(p678); npr-1(ad609) lin-15(n765ts)*. Test DNA was injected at 10–50 ng μl^{-1} for *npr-1* constructs, 20–50 ng μl^{-1} for *tax-4* constructs, and 20 ng μl^{-1} for *gcy-32::egl-2(gf)*. At least four transgenic lines were assayed for each construct; the figures show data from one representative line. Strains carrying promoter::gene transgenes were examined by DIC and fluorescence microscopy to confirm GFP expression in the appropriate cells.

Behavioural assays

Aggregation and bordering behaviours were quantified as described previously², except that assays used 80 animals and the behaviours were scored after 1 h. Locomotion speeds were quantified using DIAS software². Animals transgenic for the *hs::npr-1::gfp* construct were incubated at 33 °C for 30 min, returned to 20 °C, and aggregation and bordering were quantified over a 4-h period and compared to non-heat-shocked control transgenic animals, or to heat-shocked, non-transgenic *npr-1(ad609)* animals.

To test single animals for social feeding, 1 to 2 individuals were placed with 80 *dpy-20(e1282ts)*; *npr-1(ad609)* animals and allowed to aggregate for 1 h. *dpy-20; npr-1(ad609)* animals aggregate as strongly as *npr-1(ad609)* animals, but can be distinguished by their dumpy phenotype. Test animals were scored for whether they fed alone or in a group with *dpy-20; npr-1* animals.

The statistical significance of population and single animal behavioural assays was determined using the two-tailed *t*-test and the Mann–Whitney rank sum test, respectively.

Molecular biology

npr-1::gfp was made by inserting exons encoding GFP (A. Fire, personal communication) into a unique intronic *MluI* site of the pM2 clone². This resulted in in-frame insertion of GFP at 24 amino acids after the seventh transmembrane domain of NPR-1.

Heterologous promoter::*npr-1::gfp* constructs were made in the expression vector pPD49.26 (A. Fire, personal communication) using a 3-kilobase *npr-1* genomic DNA fragment (with GFP inserted in the *MluI* site as above) and PCR-amplified promoter fragments. The *hs::npr-1::gfp* construct was made using pPD49.78 (A. Fire, personal communication).

tax-4 and *egl-2(gf)* expression constructs were generated using the Gateway system (Invitrogen)²⁷. A library of destination vectors was created to target expression of genes of interest to specific *C. elegans* cells. The promoters used (with neuronal expression in parentheses) were: *gcy-32* (AQR, PQR, URX); *flp-8* (AWA, URX, PVM); *gcy-8* (AFD); *gcy-33* (BAG); and *odr-4* (AWA, AWB, AWC, ADE, ASG, ASH, ASI, ASJ, ASK, ADL, PHA,

PHB)^{6,8,28}. Entry vectors were designed to place the gene of interest in an artificial operon with *gfp*. Further details of plasmid construction are available on request.

Dye filling

Dil staining was used to help identify fluorescent neurons in animals that were transgenic for *npr-1::gfp*. Animals were incubated in 10 µg ml⁻¹ Dil (Molecular Probes) in M9 Medium for 30 min before examination by confocal microscopy.

Received 24 June; accepted 30 September 2002; doi:10.1038/nature01170.

1. Hodgkin, J. & Doniach, T. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149–164 (1997).
2. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**, 679–689 (1998).
3. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
4. Wang, Q. & Wadsworth, W. G. The C domain of NETRIN UNC-6 silences calcium/calmodulin-dependent protein kinase- and diacylglycerol-dependent axon branching in *Caenorhabditis elegans*. *J. Neurosci.* **22**, 2274–2282 (2002).
5. Sengupta, P., Chou, J. H. & Bargmann, C. I. *odr-10* encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell* **84**, 899–909 (1996).
6. Dwyer, N. D., Troemel, E. R., Sengupta, P. & Bargmann, C. I. Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell* **93**, 455–466 (1998).
7. Stringham, E. G., Dixon, D. K., Jones, D. & Candido, E. P. Temporal and spatial expression patterns of the small heat shock (*hsp16*) genes in transgenic *Caenorhabditis elegans*. *Mol. Biol. Cell* **3**, 221–233 (1992).
8. Yu, S., Avery, L., Baude, E. & Garbers, D. L. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc. Natl Acad. Sci. USA* **94**, 3384–3387 (1997).
9. Weinschenk, D., Wei, A., Salkoff, L. & Thomas, J. H. Block of an *ether-a-go-go*-like K⁺ channel by imipramine rescues *egl-2* excitation defects in *Caenorhabditis elegans*. *J. Neurosci.* **19**, 9831–9840 (1999).
10. Mori, I. Genetics of chemotaxis and thermotaxis in the nematode *Caenorhabditis elegans*. *Annu. Rev. Genet.* **33**, 399–422 (1999).
11. Finn, J. T., Grunwald, M. E. & Yau, K. W. Cyclic nucleotide-gated ion channels: an extended family with diverse functions. *Annu. Rev. Physiol.* **58**, 395–426 (1996).
12. Coburn, C. M. & Bargmann, C. I. A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*. *Neuron* **17**, 695–706 (1996).
13. Komatsu, H. *et al.* Functional reconstitution of a heteromeric cyclic nucleotide-gated channel of *Caenorhabditis elegans* in cultured cells. *Brain Res.* **821**, 160–168 (1999).
14. Komatsu, H., Mori, I., Rhee, J. S., Akaike, N. & Ohshima, Y. Mutations in a cyclic nucleotide-gated channel lead to abnormal thermosensation and chemosensation in *C. elegans*. *Neuron* **17**, 707–718 (1996).
15. Osborne, K. A. *et al.* Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* **277**, 834–836 (1997).
16. Ben-Shahar, Y., Robichon, A., Sokolowski, M. B. & Robinson, G. E. Influence of gene action across different time scales on behavior. *Science* **296**, 741–744 (2002).
17. Marsh, D. J., Baraban, S. C., Hollopeter, G. & Palmiter, R. D. Role of the Y5 neuropeptide Y receptor in limbic seizures. *Proc. Natl Acad. Sci. USA* **96**, 13518–13523 (1999).
18. Cowley, M. A. *et al.* Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* **411**, 480–484 (2001).
19. Lee, D. L. & Atkinson, H. J. *Physiology of Nematodes* (Macmillan, London, 1976).
20. Li, C., Nelson, L., Kyuhung, K., Nathoo, A. & Hart, A. C. Neuropeptide gene families in the nematode *Caenorhabditis elegans*. *Ann. NY Acad. Sci.* **897**, 239–252 (1999).
21. Schwartz, M. W., Woods, S. C., Porte, D. Jr, Seeley, R. J. & Baskin, D. G. Central nervous system control of food intake. *Nature* **404**, 661–671 (2000).
22. Robinson, G. E., Fahrback, S. E. & Winston, M. L. Insect societies and the molecular biology of social behavior. *Bioessays* **19**, 1099–1108 (1997).
23. Giraldeau, L.-A. & Caraco, T. *Social Foraging Theory* (Princeton Univ. Press, Princeton, New Jersey, 2000).
24. Sulston, J. & Hodgkin, J. *The Nematode Caenorhabditis elegans* (ed. Wood, W. B.) 587–606 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1988).
25. Clark, S. G., Lu, X. & Horvitz, H. R. The *Caenorhabditis elegans* locus *lin-15*, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. *Genetics* **137**, 987–997 (1994).
26. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
27. Walhout, A. J. *et al.* GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFs. *Methods Enzymol.* **328**, 575–592 (2000).
28. Li, C., Kim, K. & Nelson, L. S. FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *Brain Res.* **848**, 26–34 (1999).

Supplementary Information accompanies the paper on Nature's website (♦ <http://www.nature.com/nature>).

Acknowledgements We thank C. Bargmann, A. Fire, R. Horvitz, C. Li, I. Mori, A. Wei and S. Yu for sharing unpublished data, constructs and strains; T. Stiernagle and the *Caenorhabditis* Genetics Centre for strains; and G. Brown, N. Hopper, N. Tremain and C. Bargmann for suggestions. Some of this work was done while M.d.B. was in the laboratory of C. Bargmann. We are grateful to B. Amos and S. Reichelt for assistance with confocal microscopy.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.d.B. (e-mail: debono@mrc-lmb.cam.ac.uk).

Progenitor cell maintenance requires *numb* and *numblike* during mouse neurogenesis

Petur H. Petersen^{*†}, Kaiyong Zou^{*†}, Joseph K. Hwang^{*}, Yuh Nung Jan[‡] & Weimin Zhong^{*}

^{*} Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA

[‡] Howard Hughes Medical Institute and Departments of Physiology and Biochemistry, University of California, San Francisco, California 94143, USA

[†] These authors contributed equally to this work

Neurons in most regions of the mammalian nervous system are generated over an extended period of time during development. Maintaining sufficient numbers of progenitors over the course of neurogenesis is essential to ensure that neural cells are produced in correct numbers and diverse types^{1–3}. The underlying molecular mechanisms, like those governing stem-cell self-renewal in general, remain poorly understood. We report here that mouse *numb* and *numblike* (*Nbl*)^{4–6}, two highly conserved homologues of *Drosophila numb*^{7,8}, play redundant but critical roles in maintaining neural progenitor cells during embryogenesis, by allowing their progenies to choose progenitor over neuronal fates. In *Nbl* mutant embryos also conditionally mutant for mouse *numb* in the nervous system, early neurons emerge in the expected spatial and temporal pattern, but at the expense of progenitor cells, leading to a nearly complete depletion of dividing cells shortly after the onset of neurogenesis. Our findings show that a shared molecular mechanism, with mouse *Numb* and *Nbl* as key components, governs the self-renewal of all neural progenitor cells, regardless of their lineage or regional identities.

In *Drosophila*, *Numb* is a membrane-associated signalling protein that allows two daughter cells to adopt different fates after an asymmetric division. It does this by localizing to only one half of the cell membrane in dividing precursor cells, so that it is segregated primarily to one cell^{7,8}. On the basis of studies in the developing mouse neocortex, we have postulated that mouse *Numb* segregates to, and promotes the fate of, progenitor cells in asymmetric divisions that generate a neuron and a daughter progenitor cell during mammalian neurogenesis⁴. This view, however, is controversial; others have postulated instead that vertebrate *Numb* proteins promote the neuronal fate in such divisions^{9,10}. *numb* mutant mice exhibit severe defects in cranial neural tube closure and die around embryonic day (E) 11.5, but neurogenesis abnormalities are limited^{10,11} and insufficient to resolve the controversy. From studies to be described elsewhere, we generated *Nbl* homozygous mutant mice, which are viable, fertile and exhibit no obvious phenotypes. We find low levels of *Nbl* expression in E8.5 embryos, including in neural progenitor cells. Moreover, embryos mutant for both mouse *numb* and *Nbl* die around E9.5 with more widespread defects than single mutants.

We therefore used Cre-loxP mediated gene targeting¹² to examine their function in neurogenesis. We generated a transgenic line (*NesCre8*) with Cre expression controlled by a *nestin* promoter¹³ that is active in neural progenitor cells and somites. Within the nervous system, Cre-mediated recombination¹⁴ is readily detectable in a majority of the progenitor cells at E8.5 (*n* = 7/9) and becomes nearly complete by E12.5 (Fig. 1a, and data not shown). Conditional knockout (cKO) using *NesCre8* and a floxed mouse *numb* allele¹¹ did not cause embryonic lethality or defects in neural tube closure. In fact, cKO mice are viable, fertile and indistinguishable from their wild-type littermates (data not shown). As expected, immunoblots show mouse *Numb* protein level is already greatly reduced in E9.5 conditional mutant embryos (Fig. 1b). Most of the residual mouse

Numb protein probably comes from tissues where Cre is not active.

Mouse *numb* cKO in the *Nbl* homozygous mutant background (conditional double-knockout, or cDKO), on the other hand, results in embryonic lethality. We never recovered cDKO mice postnatally, whereas those with other allelic combinations, in particular cKO in *Nbl* heterozygous background, are viable and exhibit no gross morphological or behavioural defects. cDKO embryos are indistinguishable from their littermates at E9.5 ($n = 7$), but become completely necrotic by E12.5 ($n = 4$). Those recovered at E11.5 are considerably smaller than the littermates ($n = 8$), suggesting that cDKO embryos die around this stage.

At E10.5, cDKO embryos were consistently recovered (33/183, 1/4 or 1/8 expected). They are 80–90% the size of their wild-type or single-mutant littermates, although many (21/33) are within the range of variation seen in wild-type litters. cDKO embryos are morphologically appropriate for their age with the expected somite numbers, but have significantly reduced telencephalic vesicle (Fig. 1c, right panel) and undulating spinal cord (Fig. 1e, right), the combination of which can be reliably used to identify them. Histological analysis reveals that E10.5 cDKO embryos have severe thinning of the neural tube, from the most rostral telencephalon to caudal spinal cord, including optic discs (Fig. 1d–f, right). In 82% of the mutants, the neuroepithelium is only one-quarter to one-half the thickness of that in the littermates, with frequent buckling of the surface.

E10.5 wild-type neuroepithelium consists mainly of progenitor cells that make up the wider, inner ventricular zone, with a much smaller number of neurons forming the outer mantle zone. We could detect neurons in E10.5 cDKO embryos, in a region-specific pattern similar to that in control littermates, wild-type or other allelic combinations, using two general neuronal markers, anti-HuC/D (Hu; Fig. 2a'–c', in red) or Neurofilament (NF; Fig. 2d'–g', in green) ($n = 10$). *Dll1*, a marker for newborn, migrating neurons in the ventricular zone^{15,16}, is also expressed throughout the cDKO nervous system ($n = 5$) (Fig. 2h'–l'), including the forebrain, which has few Hu- or NF-positive neurons (Fig. 2a', h'). In fact, the cDKO neuroepithelium frequently contains large patches of *Dll1*-positive cells (Fig. 2i', l'), unlike in control embryos where they are invariably discrete (Fig. 2i, l).

To assess the severity of the neural progenitor cell loss in E10.5 cDKO embryos, we first used an antibody against phospho-Histone H3 (P-H3) to identify mitotic cells. In wild-type embryos, neural progenitor cells within the neural tube undergo S phase (DNA synthesis) when their nuclei are in the outer half of the ventricular zone. The nuclei then translocate towards the ventricular surface where cells undergo mitosis. Accordingly, P-H3-positive cells form a near continuous outline of the ventricular surface (Fig. 2a–c). In the cDKO nervous system, however, there is a dramatic reduction of P-H3-positive cells (Fig. 2a'–c'). We quantified the loss in the forebrain, the hindbrain (at the level of otic vesicle), and the spinal cord (at cervical levels), which ranges from about 80% to near 100% ($n = 5$).

We next performed bromodeoxyuridine (BrdU) labelling experiments. BrdU is incorporated into DNA by S-phase cells and, therefore, the number of cells labelled during a short pulse reflects the number of cells still proliferating but not in mitosis. There is little variation among control littermates in the pattern and percentage of neural progenitor cells labelled by BrdU (Fig. 2d–g). In forebrain, hindbrain and cervical spinal cord sections, BrdU-labelled cells account for about 49, 33 and 40% of the ventricular zone cells, respectively ($n = 3$; 1-h pulse). In contrast, only a few BrdU-positive cells are present in the cDKO nervous system, indicating a loss of over 99% of the S-phase cells ($n = 14$) (Fig. 2d'–g').

To confirm that the near-absence of proliferating cells in E10.5 cDKO embryos reflects an absence of neural progenitor cells, rather than their becoming quiescent or defective in cell-cycle progression, we examined the expression of Nestin protein, a commonly used neural progenitor cell marker. Nestin-positive cells are reduced in

numbers, but not completely absent, in the E10.5 cDKO neural tube (data not shown). As Nestin is transiently detectable in newborn neurons, probably inherited from progenitor cells, we would expect a more severe reduction at E11.5. This is indeed the case ($n = 3$) (Fig. 2m', n'). More importantly, loss of Nestin expression is specific to the nervous system; expression in somites in the same cDKO embryos is not affected (Fig. 2n', arrow).

Because the results with Nestin at E10.5 were inconclusive, we performed *in situ* hybridization using *Hes5*, which marks neural progenitor cells and is essential for keeping them undifferentiated¹⁷. As expected, there is a sharp reduction of *Hes5* expression in the E10.5 cDKO neuroepithelium ($n = 6$). In the hindbrain, for example, *Hes5*-expressing cells normally occupy two-thirds the width of the neuroepithelium (Fig. 2o). In the cDKO hindbrain, however, there is only a thin layer of scattered *Hes5*-expressing cells near the ventricular surface (Fig. 2o'). Similar loss can be observed in the ventral half of the spinal cord, where *Hes5* is highly expressed in wild-type embryos (Fig. 2p, p').

We point out that the near-complete absence of neural progenitor cells, as described above, is based on the analysis of, and invariably observed in, embryos with a neural tube that is less than half the normal thickness, a group representing 82% (27/33) of the cDKO embryos we recovered at E10.5. Furthermore, BrdU-labelled cells in

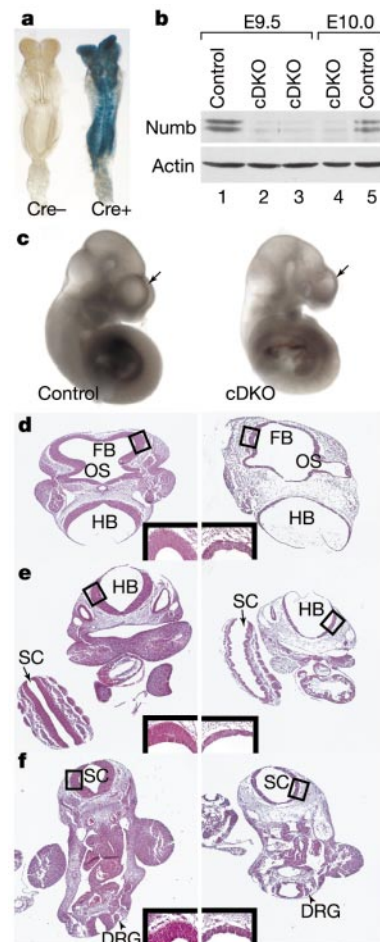


Figure 1 Conditional double knockout of mouse *numb* and *Nbl*. **a**, Cre activity at E8.5 (X-gal staining of *Rosa26* Cre-reporter¹⁴ in *NesCre8*). **b**, Mouse Numb protein level in single embryo extracts from control (1, *Nb*^{+/+}; *Nbl*^{+/+}; *NesCre8*; 5, *Nb*^{+/+}; *Nbl*^{+/+}; *NesCre8*) and cDKO (2–4) (immunoblot). **c–f**, left, An E10.5 control littermate (*Nb*^{+/+}; *Nbl*^{+/+}). **c–f**, right, An E10.5 cDKO. Sectioning of the cDKO (**c**, right) reveals severe thinning throughout the neural tube (**d–f**, right; HE staining). DRG, dorsal root ganglion; FB, forebrain; HB, hindbrain; OS, optic stalk; SC, spinal cord.

many non-neural tissues are also reduced in numbers. Whether this is a secondary effect or shows a direct requirement for mouse *numb* and *Nbl* in non-neural tissues remains to be determined.

What causes the near-absence of neural progenitor cells in E10.5 cDKO embryos? At E9.5, such embryos show no significant defects in cell proliferation (Fig. 4c' and data not shown; $n = 4$). Consistently, the overall neural tube thickness in E9.5 and E10 cDKO embryos is comparable to that in control littermates ($n = 7$ each), although the mutant neuroepithelium, particularly the ventral spinal cord, is sometimes punctuated by thinner regions. At E10, however, S-phase cells in the nervous system are reduced in numbers (Fig. 3a, b, bottom, and Fig. 4h'), particularly in regions with more active neurogenesis, indicating that the absence of progenitor cells at E10.5 is probably due to defects in self-renewal rather than smaller or defective founding populations. There is, however, significant variation in the severity of such loss among cDKO embryos ($n = 5$), probably due to the perdurance effect of the mouse Numb protein and variations in the onset of Cre expression. Consequently, E10.25 cDKO embryos show varying degrees of neural tube thinning and reduction of BrdU-labelled

cells, ranging from a near-complete loss (data not shown) to about 60% (Fig. 3c, f, bottom and inset).

Between E10 and E10.5, mostly neurons are being generated in wild-type embryos. cDKO embryos show no ectopic expression of glia markers like *Sox10* or *Olig2*^{18–20}, indicating an absence of premature gliogenesis (data not shown). We therefore examined neuron production in E10 and E10.25 cDKO embryos to ascertain whether the inability of neural progenitor cells to self-renew results from their progenies all adopting neuronal fates, which would cause an initial overproduction of neurons and, more importantly, a significant increase of their percentage within the neuroepithelium owing to a depletion of progenitor cells. Other causes, such as progenitor cells becoming quiescent or defective in cell-cycle progression, or undergoing programmed cell death (apoptosis), should affect neuron production negatively, resulting in fewer neurons, both in absolute numbers and as a percentage of the neuroepithelium.

We analysed E10.25 cDKO embryos in which neural tube thinning is not yet apparent ($n = 2$). Three lines of evidence show unambiguously that a diminished self-renewal capability among cDKO neural progenitor cells indeed results from over-differen-

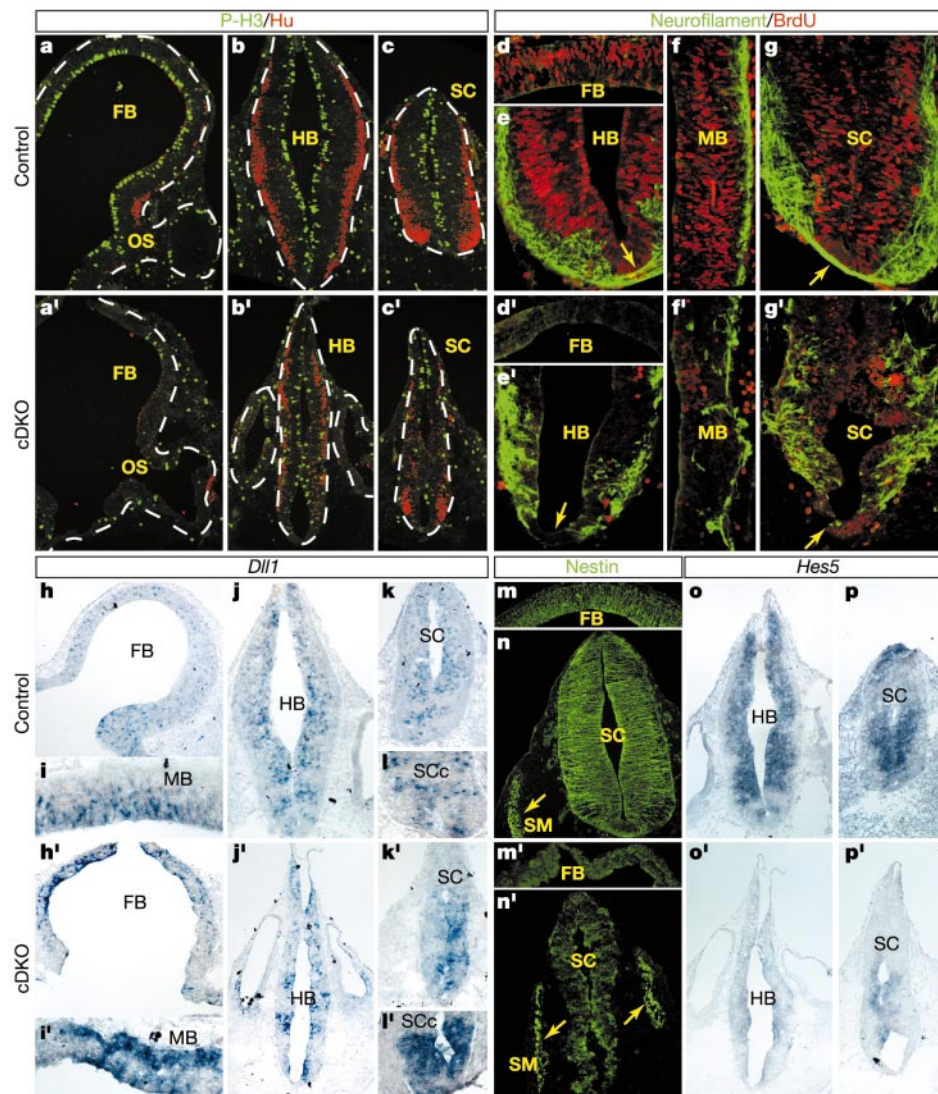


Figure 2 Effect of mouse *numb* and *Nbl* conditional double mutation on neurogenesis. **a–p**, Control littermates from E10.5 (**a–l**, **o**, **p**; *Nb*^{fl/+}; *Nbl*^{fl/+}; *NesCre8*) and E11.5 (**m**, **n**; *Nb*^{fl/+}; *Nbl*^{fl/+}). **a'–p'**, cDKO embryos from E10.5 (**a'–l'**, **o'**, **p'**) and E11.5 (**m'**, **n'**). Absence of neural progenitor cells in the cDKO is evident from the near-absence of mitotic

cells (**a'–c'**; P-H3 in green), S-phase cells (**d'–g'**; BrdU in red), and Nestin (**m'**, **n'**) and *Hes5* (**o'**, **p'**) expression. 10 × magnification except in **d–g**, **d'–g'** (25 ×) and **i**, **l**, **i'**, **l'** (20 ×). FB, forebrain; HB, hindbrain; MB, midbrain; OS, optic stalk; SC and SCc, rostral and caudal spinal cord, respectively; SM, somite.

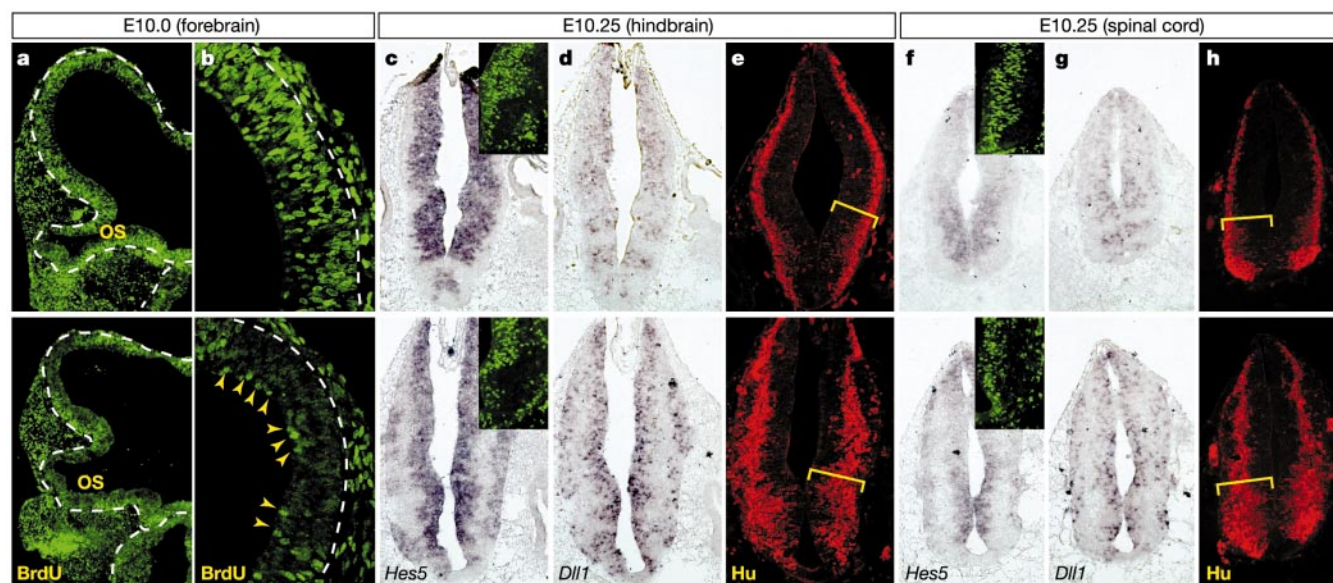


Figure 3 Effect of mouse *numb* and *Nbl* conditional double mutation on neurogenesis before neural tube thinning. **a–h**, top, Control littermates from E10.0 (**a, b**, top; *Nb*^{+/+}; *Nbl*^{+/+}; *NesCre*⁸) and E10.25 (**c–h**, top; *Nb*^{+/+}; *Nbl*^{+/+}; *NesCre*⁸). **a–h**, bottom, cDKO embryos from E10.0 (**a, b**, bottom) and E10.25 (**c–h**, bottom). Decreases in progenitor

cells (**c, f**, bottom; BrdU in green and *Hes5* expression) are accompanied by increases in *Dll1* (**d, g**, bottom) and Hu (**e, h**, bottom, in red) positive neurons. Arrowheads in **b**, bottom, point to mitotic cells labelled earlier by BrdU. OS, optic stalk. 10 × magnification except in **b**, top and bottom (40 ×).

tion of their progenies. First, reductions in the number of BrdU-labelled cells are accompanied by a similar decrease in cells expressing progenitor marker *Hes5* (Fig. 3c, f, bottom). Second, within progenitor domains marked by *Hes5*, a higher percentage of cells express *Dll1*, a marker for newborn neurons^{15,16} (Fig. 3d, g, bottom; sections are adjacent to those in Fig. 3c, f, bottom, respectively). Third and most important, there is a significant expansion of, proportional to the decrease of progenitor domains, cells expressing neuronal marker Hu (Fig. 3e, h, bottom, and Fig. 5). In these cDKO embryos, there are patches of *Hes5*-positive progenitor cells at the most lateral positions of the neuroepithelium (Fig. 3c, f, bottom). It remains to be determined whether they had migrated aberrantly, or whether their nuclei were physically prevented from translocating medially owing to neuron overproduction near the ventricular surface.

We analysed in more detail neurogenesis in the ventral spinal cord, where motor neurons emerge shortly after E9 and are continuously generated until E13.5. At E9.5 (Fig. 4a–c, a'–c'), cDKO and control embryos (*n* = 4) show little difference in Olig2 and *Isl1* expression, which marks motor progenitors and motor neurons, respectively^{21–24}. By E10, less than two cell cycles later, there are dramatic differences. At cervical levels in control embryos (Fig. 4d–f), wild-type or other allelic combinations, which are indistinguishable, *Isl1*-positive neurons colonize the lateral half of the motor domain, whereas Olig2-positive progenitor cells occupy the medial half^{21–24}. On average, only 25.6% (*n* = 6 sections from three embryos) of the Olig2-expressing cells are doubly positive for *Isl1*, representing newborn motor neurons. In contrast, *Isl1*-expressing cells frequently span the entire width of the cDKO neuroepithelium (*n* = 4/5) (Fig. 4d'–f', right). As expected, more Olig2-positive cells in the mutant co-express *Isl1*, ranging from 40.6% to as high as 68.7% (Fig. 4f', right). Even in regions where motor neurons appear to have only colonized the lateral positions, Olig2-positive cells are also absent from the ventricular zone and many, consistent with their position, co-express *Isl1* (Fig. 4d'–f', left).

Neurogenesis in the spinal cord proceeds in a rostral to caudal gradient. More caudally, there are more Olig2 single-positive cells than that at cervical levels in E10 cDKO embryos. Although the overall reduction of BrdU-labelled cells within the spinal cord is limited, it is

much more severe among Olig2-expressing cells, which sometimes show no BrdU incorporation at all (Fig. 4g'–h', right). Accordingly, in E10.25 cDKO embryos (*n* = 2) at similar caudal levels, *Isl1*-positive motor neurons span the entire width of the neuroepithelium and most of the remaining Olig2-positive cells co-express *Isl1* (Fig. 4i', j').

Although neurons are initially overproduced in cDKO embryos, there is significant increase in apoptosis among mutant neurons (Fig. 5, in red), indicating that many die shortly after birth. This is consistent with the absence of axon tracts underneath the floor plate (Fig. 2e', g', arrow), and suggests a role for mouse *numb* and *Nbl* in later events of neural development, as we have postulated⁵, although neuronal death may be a secondary effect.

It remains possible that cDKO progenitor cells adopt an abnormal developmental pathway and differentiate *en masse* before the onset of neurogenesis. However, two lines of evidence strongly suggest that they are rapidly depleted as neurogenesis progresses because their daughter cells all adopt neuronal fates instead of self-renewing after division. First, in cDKO embryos where BrdU-labelled S-phase cells are significantly reduced in numbers but not totally absent, the number of mitotic cells is comparable to that in control littermates (Fig. 3b, bottom, arrowheads, Fig. 5, P-H3, in blue). Second, whereas neuron overproduction is observed throughout the cDKO nervous system, it is more pronounced in regions where neurogenesis has been more active. In the control E10.25 forebrain, for example, only a few scattered Hu-positive neurons are present (Fig. 5a, left), indicating that neurogenesis was just underway. Neurons are overproduced in the cDKO forebrain, but represent only a small fraction of the neuroepithelium (Fig. 5a, right). On the other hand, ventral spinal cord in control littermates contains large numbers of neurons (Fig. 5e, left). Accordingly, Hu-positive neurons not only are overproduced in the cDKO but also span nearly the entire width of the neuroepithelium (Fig. 5e, right). Similar differences in neuron overproduction can be observed in other regions (Fig. 5b–d, right) and along the dorsoventral neural axis (Fig. 3h, bottom).

The findings reported here are consistent with our earlier hypothesis^{4,5}, and demonstrate unequivocally that the main function of *numb* homologues in mouse neurogenesis is to maintain progenitor cells, not promoting neuronal fates, during the initial progenitor

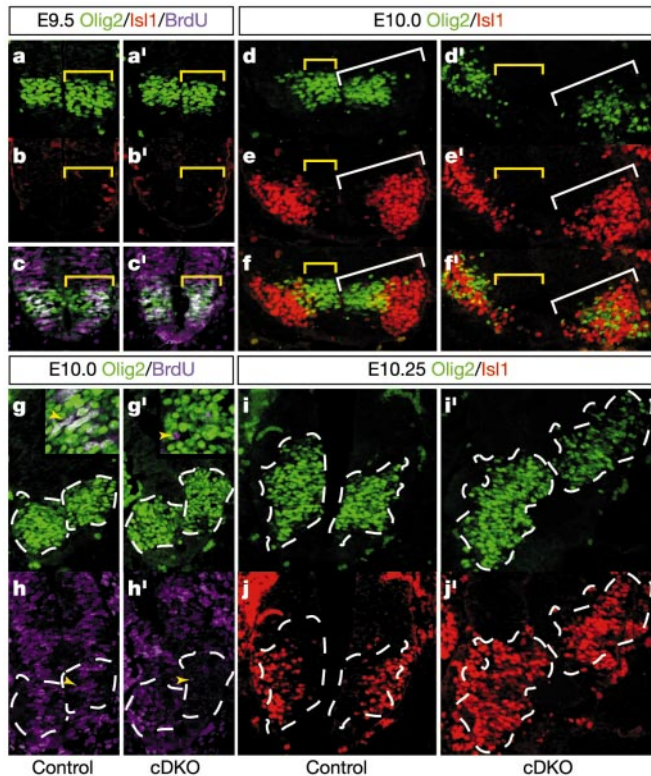


Figure 4 Motor neuron generation at the expense of progenitor cells. **a–j**, Control littermates from E9.5 (**a–c**; *Nbl*^{+/+}; *Nbl*^{+/+}), E10.0 (**d–h**; *Nbl*^{+/+}; *Nbl*^{+/+}; *NesCre8*), and E10.25 (**i, j**; *Nbl*^{+/+}; *Nbl*^{+/+}; *NesCre8*). **a'–j'**, cDKO embryos from E9.5 (**a'–c'**), E10.0 (**d'–h'**), and E10.25 (**i', j'**). Cervical (**a–f**, **a'–f'**) or caudal (**g–j**, **g'–j'**) spinal cord sections were double-labelled for Olig2 (in green) and Isl1 (in red) or BrdU (in pink). After E9.5, proliferating motor progenitor cells decrease rapidly in numbers in cDKO embryos (**g', h'**). Consistently, cDKO motor neurons often span the entire width of the neuroepithelium (**d'–f'**, **i', j'**).

versus neuronal fate decision. This is also supported by gain-of-function studies in chick showing an increase in progenitors among neuroepithelial cells over-expressing chick *Numb*⁹. A near-absence of motor and sensory neurons in mouse *numb* single mutants at E10.5 due to impaired differentiation has been reported¹⁰. This phenotype is not seen in our single or cDKO mutants and, therefore, is unlikely to be due to a loss of mouse *numb* function.

Our findings are, to our knowledge, the first direct evidence of a pan-neural program for precursor cells, regardless of their lineage or regional identities, to choose between proliferation and differentiation. There is growing evidence from direct imaging experiments that asymmetric cell division occurs during mammalian neurogenesis^{25–27}. Mouse *Numb* is asymmetrically localized to the apical membrane in dividing progenitor cells^{4,5,28}, but how *Nbl* protein is distributed in these cells is unknown, owing to low levels of expression. Therefore, although an effect on asymmetric division can account for pan-neural progenitor depletion in cDKO embryos (Fig. 5f, left), further studies are necessary to ascertain this, in particular the specific effects on multipotential neural stem cells and progenitors with more limited developmental potentials. Similarly, to ascertain if mouse *Numb* and *Nbl* act by inhibiting Notch activity like *Drosophila* *Numb*, it is necessary to determine first whether Notch signalling plays a generic role in regulating cell fate choices between proliferation and differentiation—as seen with mouse *Numb* and *Nbl*—or acts differently in different progenitor populations.

Finally, the widespread defects exhibited by mouse *numb* and *Nbl* constitutive double mutants at E9.5 raise an interesting possibility

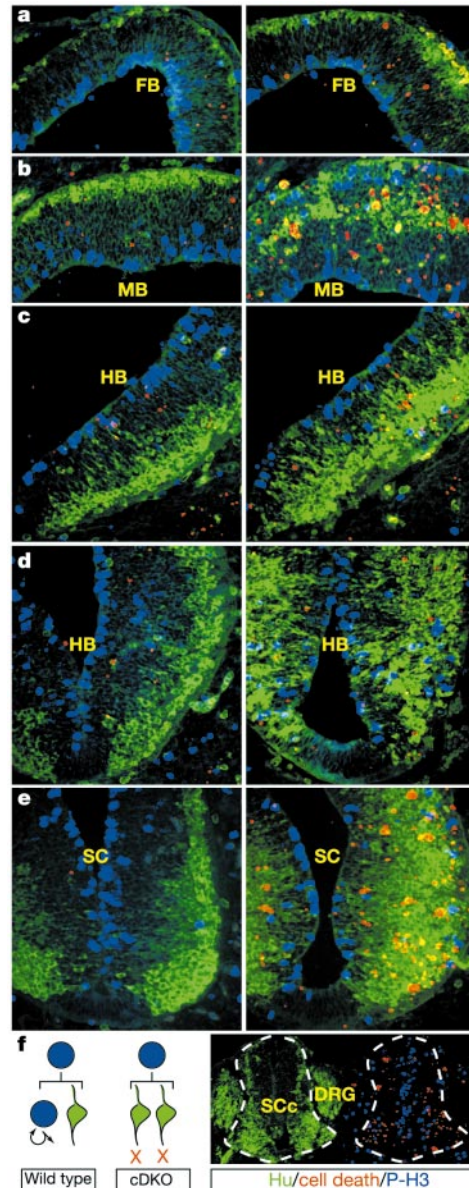


Figure 5 Neuron overproduction and death in mouse *numb* and *Nbl* conditional double-mutant embryos. Compared to the control (**a–e**, left), triple labelling reveals that increases in apoptosis (in red) in the cDKO (**a–f**, right) are largely confined to regions with Hu-positive neurons (in green). As shown schematically in **f**, left, nervous system phenotypes in cDKO mutants probably result from both daughters of neural progenitor cells (blue filled circle) adopting neuronal fates (in green) and then undergoing apoptosis (red X). Embryos (E10.25) are the same as in Fig. 3. Magnification in **f**, right, is two-fifths of that in parts **a–e**. DRG, dorsal root ganglion; FB, forebrain; HB, hindbrain; MB, midbrain; SC and SCc, rostral and caudal spinal cord, respectively. HB sections in **c**, left and right, are rostral to those in **d**, left and right, respectively.

that these two genes are involved in stem-cell self-renewal in other tissues. If mouse *Numb* and *Nbl* proteins indeed act like their *Drosophila* counterpart, they would provide an attractive molecular mechanism to integrate lineage or cell-extrinsic cues for progenies of various stem cells to choose between self-renewal and adopting appropriate differentiated fates. □

Methods

Generation of mutant mice

The *Nbl* null (–) allele, lacking 90% of the coding region, was generated using homologous recombination in ES cells and will be described elsewhere. The floxed (f) and

null (also named $\Delta 5,6$) alleles of mouse *numb* were generated previously¹¹. The *NesCre8* transgene was constructed by inserting a Cre complementary DNA with a nuclear localization signal and chick β -actin poly-adenylation site (a gift from M. Lewandoski) immediately 3' to the *nestin* transcription start site in a vector with 5.8 kilobases (kb) of upstream promoter and 3.2 kb of downstream sequence, including the first two introns¹³ (a gift from U. Lendahl). Transgenic mice were produced as previously described²⁹. cDKO embryos were generated by mating *numb*^{0/0}; *Nbl*^{-/-} males with *numb*^{+/+}; *Nbl*^{+/+}; *NesCre8* females. The latter were either hemizygous or homozygous for the Cre transgene.

Histology and embryo staining

X-gal staining, immunoblot, immunostaining, BrdU labelling, and *in situ* hybridization were as previously described with minor modifications^{4,5,11,29}. Anti-Actin, BrdU (Sigma), HuC/D (Molecular Probes), Isl1 and Nestin (Developmental Studies Hybridoma Bank) are mouse monoclonal, whereas anti-NF (*m*, 150K, Chemicon), Olig2 (a gift from H. Takebayashi) and P-H3 (Upstate Biotechnology) are rabbit polyclonal. Apoptosis was detected using a kit (Roche). Probes for *Dll1* and *Hes5* are based on mouse and rat cDNAs, respectively.

Received 18 June; accepted 23 August 2002; doi:10.1038/nature01124.

- Anderson, D. A. Stem cells and pattern formation in the nervous system: the possible versus the actual. *Neuron* **30**, 19–35 (2001).
- Edlund, T. & Jessell, T. M. Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system. *Cell* **96**, 211–224 (1999).
- McConnell, S. K. Constructing the cerebral cortex: neurogenesis and fate determination. *Neuron* **15**, 761–768 (1995).
- Zhong, W., Feder, J. N., Jiang, M.-M., Jan, L. Y. & Jan, Y. N. Asymmetric localization of a mammalian Numb homolog during mouse cortical neurogenesis. *Neuron* **17**, 43–53 (1996).
- Zhong, W., Jiang, M.-M., Weinmaster, G., Jan, L. Y. & Jan, Y. N. Differential expression of mammalian Numb, Numblike and Notch1 suggests distinct roles during mouse cortical neurogenesis. *Development* **124**, 1887–1897 (1997).
- Verdi, J. M. *et al.* Mammalian NUMB is an evolutionarily conserved signaling adapter protein that specifies cell fate. *Curr. Biol.* **6**, 113–145 (1996).
- Rhyu, M. S., Jan, L. Y. & Jan, Y. N. Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. *Cell* **76**, 477–491 (1994).
- Spana, E. P., Koczynski, C., Goodman, C. S. & Doe, C. Q. Asymmetric localization of numb autonomously determines sibling neuron identity in the *Drosophila* CNS. *Development* **121**, 3489–3494 (1995).
- Wakamatsu, Y., Maynard, T. M., Jones, S. U. & Weston, J. A. NUMB localizes in the basal cortex of mitotic avian neuroepithelial cells and modulates neuronal differentiation by binding to NOTCH-1. *Neuron* **23**, 71–81 (1999).
- Zilian, O. *et al.* Multiple roles of mouse *numb* in tuning developmental cell fates. *Curr. Biol.* **11**, 494–501 (2001).
- Zhong, W. *et al.* Mouse *numb* is an essential gene involved in cortical neurogenesis. *Proc. Natl Acad. Sci. USA* **97**, 6844–6849 (2000).
- Gu, M., Marth, J. D., Orban, P. C., Mossman, H. & Rajewsky, K. Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103–106 (1994).
- Zimmerman, L. *et al.* Independent regulatory elements in the *nestin* gene direct transgene expression to neural stem cells or muscle precursors. *Neuron* **12**, 11–24 (1994).
- Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genet.* **21**, 70–71 (1999).
- Dunwoodie, S. L., Henrique, D., Harrison, S. M. & Beddington, R. S. P. Mouse *Dll3*: a novel divergent *Delta* gene which may complement the function of other *Delta* homologues during early pattern formation in the mouse embryo. *Development* **124**, 3065–3076 (1997).
- Campos, L. S., Duarte, A. J., Branco, T. & Henrique, D. *mDll1* and *mDll3* expression in the developing mouse brain: role in the establishment of the early cortex. *J. Neurosci. Res.* **64**, 590–598 (2001).
- Ohtsuka, T. *et al.* *Hes1* and *Hes5* as Notch effectors in mammalian neuronal differentiation. *EMBO J.* **18**, 2196–2207 (1999).
- Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I. & Wegner, M. Sox10, a novel transcriptional modulator in glial cells. *J. Neurosci.* **18**, 237–250 (1998).
- Lu, Q. R. *et al.* Sonic hedgehog-regulated oligodendrocyte lineage genes encoding bHLH proteins in the mammalian central nervous system. *Neuron* **25**, 317–329 (2000).
- Zhou, Q., Wang, S. & Anderson, D. J. Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors. *Neuron* **25**, 331–343 (2000).
- Takebayashi, H. *et al.* Dynamic expression of basic helix-loop-helix Olig family members: implication of Olig2 in neuron and oligodendrocyte differentiation and identification of a new member, Olig3. *Mech. Dev.* **99**, 143–148 (2000).
- Mizuguchi, R. *et al.* Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron* **31**, 757–771 (2001).
- Novitsch, B. G., Chen, A. I. & Jessell, T. M. Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* **31**, 773–789 (2001).
- Pfaff, S. L., Mendelsohn, M., Stewart, C. L., Edlund, T. & Jessell, T. M. Requirement for LIM homeobox gene *Isl1* in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. *Cell* **84**, 309–320 (1996).
- Chenn, A. & McConnell, S. K. Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis. *Cell* **82**, 631–641 (1995).
- Qian, X., Goderie, S. K., Shen, Q., Stern, J. H. & Temple, S. Intrinsic programs of patterned cell lineages in isolated vertebrate CNS ventricular zone cells. *Development* **125**, 3143–3152 (1998).
- Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714–720 (2001).
- Cayouette, M., Whitmore, A. V., Jeffery, G. & Raff, M. Asymmetric segregation of Numb in retinal development and the influence of the pigmented epithelium. *J. Neurosci.* **21**, 5643–5651 (2001).
- Zhong, W., Mirkovitch, J. & Darnell, J. E. Jr Tissue specific regulation of mouse *hepatocyte nuclear factor 4* expression. *Mol. Cell. Biol.* **14**, 7276–7284 (1994).

Acknowledgements We thank members of the Zhong laboratory for discussions, C. Jacobs and H. Keshishian for comments, K. Jakobsdottir for support (P.H.P.), J. Zhang for technical assistance, Developmental Studies Hybridoma Bank, D. Anderson, R. Kageyama, U. Lendahl, M. Lewandoski, M. Nakafuku, H. Takebayashi and M. Wegner for reagents. K.Z. is supported by a Yale Anderson fellowship. Y.N.J. is a HHMI investigator. This work was supported by a Yale Hellman Family fellowship and grants from March of Dimes and NINDS to W.Z. and from NIMH to the Silvio Conte Center for Neuroscience Research at UCSF.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to W.Z. (e-mail: weimin.zhong@yale.edu).

N-CoR controls differentiation of neural stem cells into astrocytes

Ola Hermanson*†, Kristen Jepsen* & Michael G. Rosenfeld*

* Howard Hughes Medical Institute, Department of Molecular Medicine, University of California, San Diego, School of Medicine, 9500 Gilman Drive, Room 345, La Jolla, California 92093-0648, USA

Understanding the gene programmes that regulate maintenance and differentiation of neural stem cells is a central question in stem cell biology. Virtually all neural stem cells maintain an undifferentiated state and the capacity to self-renew in response to fibroblast growth factor-2 (FGF2)^{1–5}. Here we report that a repressor of transcription, the nuclear receptor co-repressor (N-CoR), is a principal regulator in neural stem cells, as FGF2-treated embryonic cortical progenitors from *N-CoR* gene-disrupted mice display impaired self-renewal and spontaneous differentiation into astroglia-like cells. Stimulation of wild-type neural stem cells with ciliary neurotrophic factor (CNTF), a differentiation-inducing cytokine³, results in phosphatidylinositol-3-OH kinase/Akt1 kinase-dependent phosphorylation of N-CoR, and causes a temporally correlated redistribution of N-CoR to the cytoplasm. We find that this is a critical strategy for cytokine-induced astroglia differentiation and lineage-characteristic gene expression. Recruitment of protein phosphatase-1 to a specific binding site on N-CoR exerts a reciprocal effect on the cellular localization of N-CoR. We propose that repression by N-CoR, modulated by opposing enzymatic activities, is a critical mechanism in neural stem cells that underlies the inhibition of glial differentiation.

N-CoR was initially defined as a regulator of nuclear receptor-mediated repression⁶, but has been shown to interact with and act as a co-repressor for many other transcription factors including CBF1 (suppressor of hairless, Lag-1, CSL, RBP-Jk)⁷, which acts downstream of the transmembrane receptor notch⁸. Notch signalling has long been implicated in the inhibition of neuronal differentiation, and FGF2-mediated maintenance of notch expression and responsiveness may be one characteristic of the neural stem cell^{8–10}; however, *in vivo* evidence for a role of N-CoR in this context is lacking. N-CoR protein levels are high in undifferentiated, proliferating neural progenitors located in the ventricular zones of the embryonic brain¹¹ (Fig. 1a, and data not shown). Previous evaluation of *N-CoR* gene-deleted mice (*N-CoR*^{-/-}), which generally die before embryonic day 16 (E16)¹¹, revealed that neural induction and early development proceed grossly normal through to E12.5 in the absence of N-CoR. However, the brains of *N-CoR*^{-/-} mice show progressively more severe aberrations beginning at E14.5, including

† Present address: Department of Cell and Molecular Biology, Medical Nobel Institute, Karolinska Institute, SE-171 77 Stockholm, Sweden

a decrease in size of specific brain regions and decreased forebrain expression of nestin, an intermediate filament expressed in neural progenitors and radial glial cells¹¹. These data suggest that N-CoR may have an important role in nestin-expressing neural stem cells.

To investigate more conclusively the role of N-CoR in neural stem cells, we isolated cortical progenitor cells from individual mouse embryos at E13, a time which preceded any detectable differences in progenitor cells between *N-CoR*^{+/+} and *N-CoR*^{-/-} mice, as assessed by markers and morphology¹¹ (Fig. 1, and data not shown). In the presence of FGF2, progenitor cells from wild-type embryos proliferated as assessed by cell division and the expression of proliferative markers (Ki67), and formed colonies of self-renewing, nestin-positive neural stem cells (Fig. 1b) that could be passaged and differentiated into neurons and glia. In contrast, although cortical progenitors derived from E13 *N-CoR*^{-/-} mice exhibited identical plating efficiency, morphology, and initial nestin expression as those derived from wild-type mice, they did not form colonies and either completely failed to proliferate or underwent

one or two cell cycles (Fig. 1b). After 4–6 days in culture, many *N-CoR*^{-/-} cells displayed clear morphological signs of differentiation (Fig. 1b), and these events occurred even in the presence of higher concentrations of FGF2 (up to 200 ng ml⁻¹).

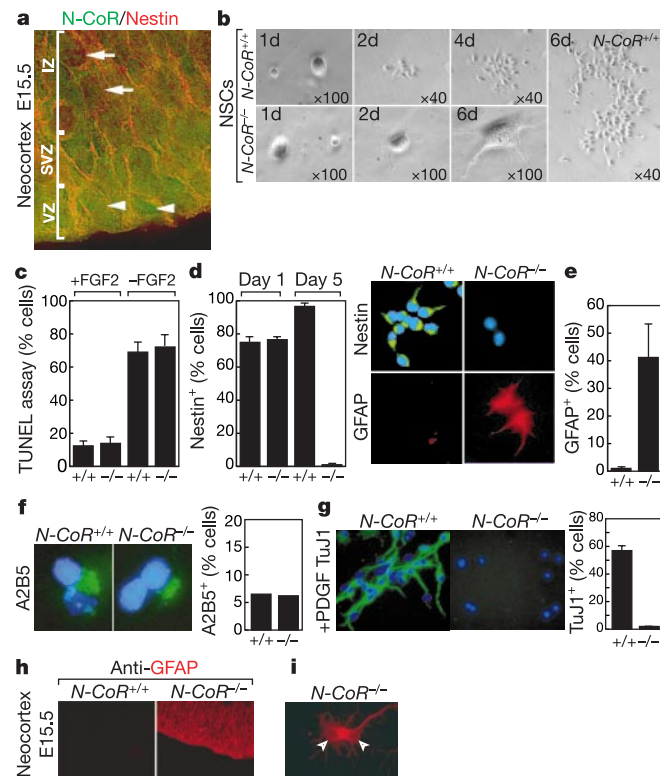


Figure 1 N-CoR is a regulator of neural stem cell state. **a**, N-CoR (green) and nestin (red) immunolabelling in neocortex at E15.5. IZ, intermediate zone; SVZ, subventricular zone; VZ: ventricular zone. Arrowheads indicate N-CoR labelling in VZ. Arrows indicate lack of nuclear N-CoR labelling. **b**, Nomarski micrographs of living *N-CoR*^{+/+} and *N-CoR*^{-/-} neural stem cells (NSCs) after 1–6 days of culture in the presence of FGF2. **c**, Numbers of apoptotic cells in *N-CoR*^{+/+} and *N-CoR*^{-/-} neural stem cells as assessed by TUNEL assay. **d**, Numbers of nestin-expressing cells in *N-CoR*^{+/+} and *N-CoR*^{-/-} cells after 1 and 5 days in culture. The right panel shows immunofluorescent labelling of nestin (green) and GFAP (red) expression in *N-CoR*^{+/+} and *N-CoR*^{-/-} cells after 5 days in the presence of FGF2. Nuclear 4,6-diamidino-2-phenylindole (DAPI) staining is indicated (blue). **e**, Numbers of GFAP-expressing, astroglia-like cells in *N-CoR*^{+/+} and *N-CoR*^{-/-} cells after 5 days of FGF2 treatment. **f**, Immunofluorescent labelling and numbers of A2B5-positive (green) cells in *N-CoR*^{+/+} and *N-CoR*^{-/-} cells. Nuclear DAPI staining is indicated (blue). **g**, Immunofluorescent labelling and numbers of the total increase of TuJ1-positive (green) cells in *N-CoR*^{+/+} and *N-CoR*^{-/-} cultures after 5 days of PDGF treatment. **h**, Immunohistochemical detection of GFAP in E15.5 *N-CoR*^{+/+} and *N-CoR*^{-/-} neocortex. **i**, Binucleate GFAP-expressing (red) *N-CoR*^{-/-} cell. Arrowheads indicate cell nuclei.

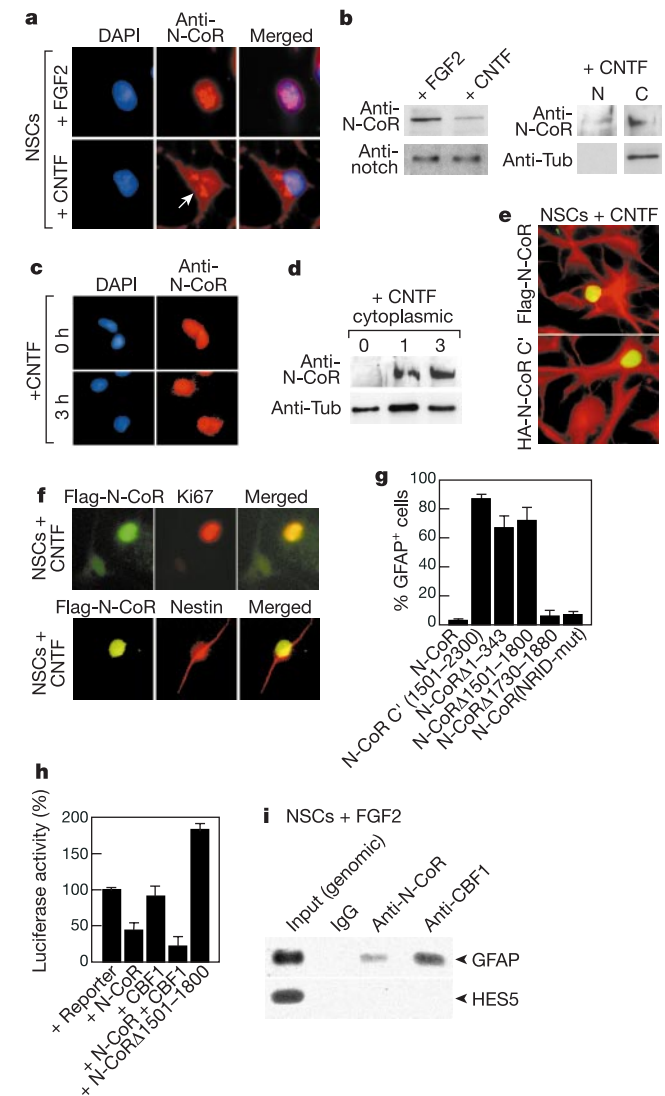


Figure 2 The subcellular localization of N-CoR is altered by CNTF, and nuclear N-CoR represses CNTF-mediated astroglia differentiation and GFAP expression. **a**, N-CoR immunoreactivity (red) and nuclear DAPI staining (blue) in the presence of FGF2 or after 24 h of CNTF treatment. **b**, Immunoblotting of whole-cell neural stem cell extracts using antibodies against N-CoR or the transmembrane receptor notch. Arrow indicates cytoplasmic N-CoR labelling. N, nuclear; C, cytoplasmic. **c**, N-CoR immunoreactivity as assessed by immunohistochemistry after 0 and 3 h of CNTF treatment. **d**, Immunoblotting of cytoplasmic fractions of neural stem cells after the indicated times following CNTF treatment using anti-N-CoR and anti- β -tubulin antibodies. **e**, Immunohistochemistry using anti-Flag or anti-HA antibodies to detect transfected Flag-tagged N-CoR or HA-tagged N-CoR C' (green) in CNTF-treated, GFAP-positive (red) cell cultures. NSCs, neural stem cells. **f**, Immunohistochemistry using anti-Flag (green) antibody to detect transfected Flag-tagged N-CoR in Ki67 (red) or nestin (red) positive neural stem cells. **g**, Proportion of transfected neural stem cells, compared with control transfections, exhibiting GFAP labelling and astroglia-like morphology after 48 h of CNTF treatment as assessed by quantification of double-labelled cells. **h**, Relative luciferase activity in CNTF-treated neural stem cells transfected with the CNTF-responsive repressor domain of the proximal *GFAP* promoter coupled to luciferase. **i**, ChIP analysis of FGF2-treated neural stem cells using anti-N-CoR, anti-CBF1, or pre-immune IgG and primers flanking the CBF1 binding site in the *GFAP* promoter. Primers against an unrelated gene exon (HES5) serve as negative control.

When equal numbers of *N-CoR*^{+/+} and *N-CoR*^{-/-} neural stem cells were cultured together, *N-CoR*^{-/-} cells were unable to proliferate and appeared differentiated after 5 days in culture, demonstrating that the defect was cell-autonomous (data not shown). In the presence of FGF2, there were no signs of increased cell death in *N-CoR*^{-/-} cultures compared with *N-CoR*^{+/+} cultures as assessed by a TdT-mediated dUTP nick end labelling (TUNEL) assay (Fig. 1c), consistent with the lack of increased cortical apoptosis in *N-CoR*^{-/-} mice¹¹. Both *N-CoR*^{+/+} and *N-CoR*^{-/-} cells cultured in the absence of FGF2 showed markedly increased apoptotic cell death after 2 days (Fig. 1c), suggesting that N-CoR is not required for the reported anti-apoptotic effect of FGF2 (ref. 12). Thus, the phenotype of *N-CoR*^{-/-} cells is due to an intrinsic inability to maintain the FGF2-mediated undifferentiated and proliferative state.

In contrast to wild-type neural stem cells, a large number (>40%, Fig. 1b–d) of *N-CoR*^{-/-} cells exhibited an astroglia-like morphology after 4–6 days in culture, were nestin negative, and expressed the intermediate filament protein glia fibrillary acidic protein (GFAP), a widely used marker of astroglia differentiation (Fig. 1d, e). To determine whether the increased GFAP expression was a result of differentiation of multipotent progenitors or was due to an increase in numbers of glia-restricted progenitors in *N-CoR*^{-/-} cultures, we investigated the expression of A2B5, a marker for a subset of glia-restricted progenitors¹³. There was no difference in the numbers of cells expressing A2B5 in neural stem cells from either *N-CoR*^{+/+} or *N-CoR*^{-/-} mice 1 day after plating (Fig. 1f). Our experiments cannot exclude the possibility of enrichment in E13 *N-CoR*^{-/-} cells of an uncharacterized, FGF2-resistant glia-restricted progenitor. To investigate whether the GFAP-expressing cells in *N-CoR*^{-/-} cultures were terminally differentiated or belonged to a class of GFAP-expressing cells with neurogenic

potential¹, we treated cell cultures with platelet-derived growth factor (PDGF), which increases neuronal differentiation of neural stem cells after initial FGF2 expansion. Treatment of neural stem cells with PDGF³ resulted in a significant increase in the number of neuronal (TuJ1-positive) cells in *N-CoR*^{+/+} ($4.2 \pm 2.3\%$ without PDGF compared with $61.6 \pm 3.6\%$ with PDGF) but not *N-CoR*^{-/-} cultures ($5.3 \pm 2.0\%$ without PDGF compared with $7.1 \pm 1.2\%$ with PDGF; Fig. 1g). As the percentage of TuJ1-positive cells before PDGF treatment is similar in both *N-CoR*^{+/+} and *N-CoR*^{-/-} cultures, our data indicate that, in contrast to the spontaneous differentiation along the glial pathway, neuronal differentiation is not induced by the absence of N-CoR. The present results suggest that the increased number of GFAP-positive cells in *N-CoR*^{-/-} cultures is the result of terminal differentiation of presumptive neural stem cells that behaved normally at E13, but subsequently failed to respond to FGF2.

In agreement with the increased GFAP expression in *N-CoR*^{-/-} cells, >70% of brains from *N-CoR*^{-/-} mice showed prominent GFAP expression as early as E14.5–15.5, a time at which wild-type brains virtually never express GFAP^{11,14} (Fig. 1h). We also noted that many ($43 \pm 4\%$) GFAP-expressing cells in the *N-CoR*^{-/-} cultures were binucleate (compared with <1% of cells in wild-type cultures). This observation suggests that *N-CoR*^{-/-} cells can progress through one cell cycle, but fail to undergo cytokinesis (Fig. 1i). Such defects have previously been reported in conjunction with overexpression and impaired phosphorylation of intermediate filaments, including GFAP¹⁵.

To determine whether regulation of N-CoR was a principal component of cell type determination, cells were treated with CNTF, which initiates a rapid (48 h) and efficient (>90%) differentiation of rat and mouse neural stem cells into GFAP-expressing astrocytic cells³. Whereas endogenous N-CoR in FGF2-treated neural stem cell cultures was predominantly nuclear, we observed that CNTF treatment resulted in an increase of cytoplasmic N-CoR and a concomitant decrease of N-CoR in the nucleus, confirmed by subcellular fractionation (Fig. 2a, b). Immunoblotting of whole-cell lysates showed that total N-CoR levels decreased, whereas notch levels remained unchanged, suggesting that translocation of N-CoR to the cytoplasm may precede degradation (Fig. 2b). Indeed, time course analysis of N-CoR localization revealed a rapid increase in cytoplasmic N-CoR, easily detectable just 1–3 h after CNTF administration, and a subsequent decrease in nuclear N-CoR (Fig. 2c, d, and data not shown). Introduction of wild-type N-CoR expression vectors into rat neural stem cells by lipofection permitted nuclear N-CoR protein levels to be maintained for 72 h even in the presence of CNTF, and resulted in cells that were indistinguishable from undifferentiated neural stem cells (Fig. 2e, f, and data not shown). These results suggest that overexpression of N-CoR inhibits the astrocytic differentiation and GFAP expression that normally occurs in response to CNTF stimulation (Fig. 2e).

To determine the domains of N-CoR required for repression of the CNTF-mediated astroglia differentiation, we evaluated a series of N-CoR deletion mutants. Constructs with deletion of amino acids 1830–1945, a repressor domain that interacts with the co-repressors mSin3 A/B, or mutation of the nuclear receptor interaction domain (N-CoR(NRID-mut))¹⁶, were as efficient as wild type N-CoR at inhibiting CNTF-mediated differentiation (Fig. 2g). However, N-CoR C-terminal (amino acids 1501–2300) as well as constructs in which either the repressor-domain-containing amino terminus (amino acids 1–343) or the CBF1-interacting domain (amino acids 1501–1800) were deleted, did not repress CNTF-mediated differentiation or GFAP expression (Fig. 2g). Overexpression of a series of other transcriptional co-repressors did not affect CNTF-induced astrocytic differentiation significantly (data not shown), indicating the specificity of the effects of N-CoR.

Previous analyses of the GFAP promoter revealed that the proximal promoter contains a *cis* element required for activation

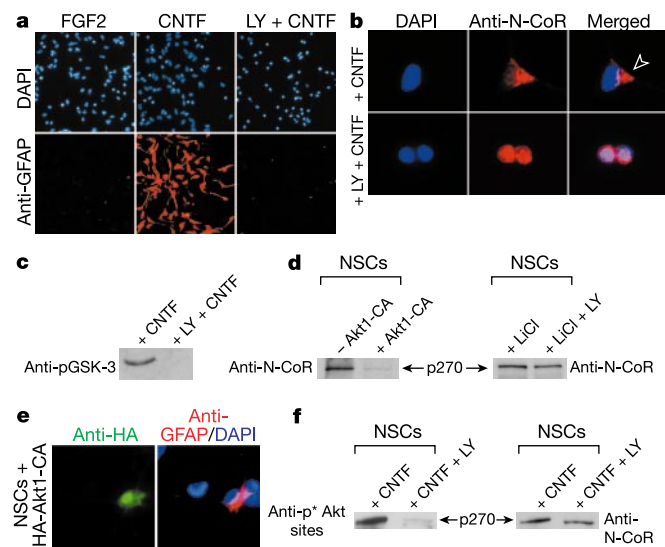


Figure 3 Akt1 kinase activation in neural stem cells is required for CNTF-mediated astroglia differentiation, GFAP expression, and re-localization of N-CoR. **a**, Micrographs of neural stem cells treated for 24 h as indicated. LY, LY294002. **b**, Immunolabelling with anti-N-CoR antibody in cells after 24 h of CNTF treatment with or without LY pre-treatment. Arrowhead indicates cytoplasmic N-CoR labelling. **c**, Immunoprecipitation of Akt1 kinase from neural stem cells treated as indicated, followed by immunoblotting using an antibody against phosphorylated GSK-3 α/β . **d**, Western blotting of neural stem cell extracts demonstrating the levels of endogenous N-CoR in neural stem cells 48 h after transfection with constitutively active Akt1 kinase (Akt1-CA) versus lithium chloride (LiCl) treatment. **e**, Immunohistochemical labelling of neural stem cells for transfected HA-tagged Akt1-CA (green), GFAP (red), and nuclear DAPI (blue) after 48 h in the presence of FGF2. **f**, Immunoblotting of CNTF-treated neural stem cell extracts in the absence or presence of LY using an antibody against phosphorylated Akt1 kinase target sites. The blots were stripped and re-probed with an anti-N-CoR antibody.

by CNTF^{17–19}, as well as a putative repressor domain which we found harbours a conserved CBF1 consensus-binding site^{17–19}. CNTF-stimulated activation of the proximal rat *GFAP* promoter (–384 base pairs (bp) to +13 bp) was reduced 60% by N-CoR overexpression (Fig. 2h). Overexpression of CBF1 alone exerted no repression, and overexpression of both N-CoR and CBF1 showed no significant increase in repression (Fig. 2h). Deletion of the CBF1-interacting domain of N-CoR (amino acids 1501–1800 (ref. 7)) rendered N-CoR unable to repress CNTF-stimulated activation of the *GFAP* promoter (Fig. 2h). Chromatin immunoprecipitation (ChIP) experiments using FGF2-treated rat neural stem cell cultures revealed that both CBF1 and N-CoR could bind to the repressor region of the *GFAP* promoter, but not to a control region in HES5 (Fig. 2i). Thus, N-CoR is probably recruited by CBF1 and acts directly on the proximal *GFAP* promoter to repress transcription.

To investigate which pathway/s downstream of CNTF regulates the subcellular localization of N-CoR and astroglia differentiation, we used a series of inhibitors of various intracellular signalling pathways suggested to function downstream of CNTF. Administration of the phosphatidylinositol-3-OH kinase (PI(3)K) 110

inhibitor LY294002 (LY) before CNTF addition resulted in a complete inhibition of CNTF-mediated astrocytic differentiation, *GFAP* expression, and N-CoR translocation (Fig. 3a, b). Because PI(3)K is known to activate Akt kinase/protein kinase B²⁰, endogenous Akt kinase was immunoprecipitated from rat neural stem cells 20 min after CNTF stimulation. We found that it was capable of phosphorylating a peptide corresponding to the N terminus of glycogen synthase kinase (GSK)–3 β , a well-established Akt kinase substrate, and that this effect was blocked by LY (Fig. 3c). Co-transfection of plasmids expressing N-CoR and constitutively active (CA) Akt1 kinase (Akt1-CA; haemagglutinin (HA)-tagged Akt1(T308D/S473D)) resulted in a marked decrease in N-CoR expression after 48 h (Fig. 3d), and treatment with the GSK–3 β inhibitor lithium chloride (LiCl) showed that these effects by Akt1-CA were not secondary to an Akt1-mediated inhibition of GSK–3 β activity (Fig. 3d). Overexpression of Akt1-CA, which was constitutively nuclear in neural stem cells, induced rapid signs of differentiation (including morphological changes with extended filopodia) and increased cell soma (resembling astroglia morphology) in 100% of the transfected neural stem cells within 48 h (Fig. 3e). The Akt1-

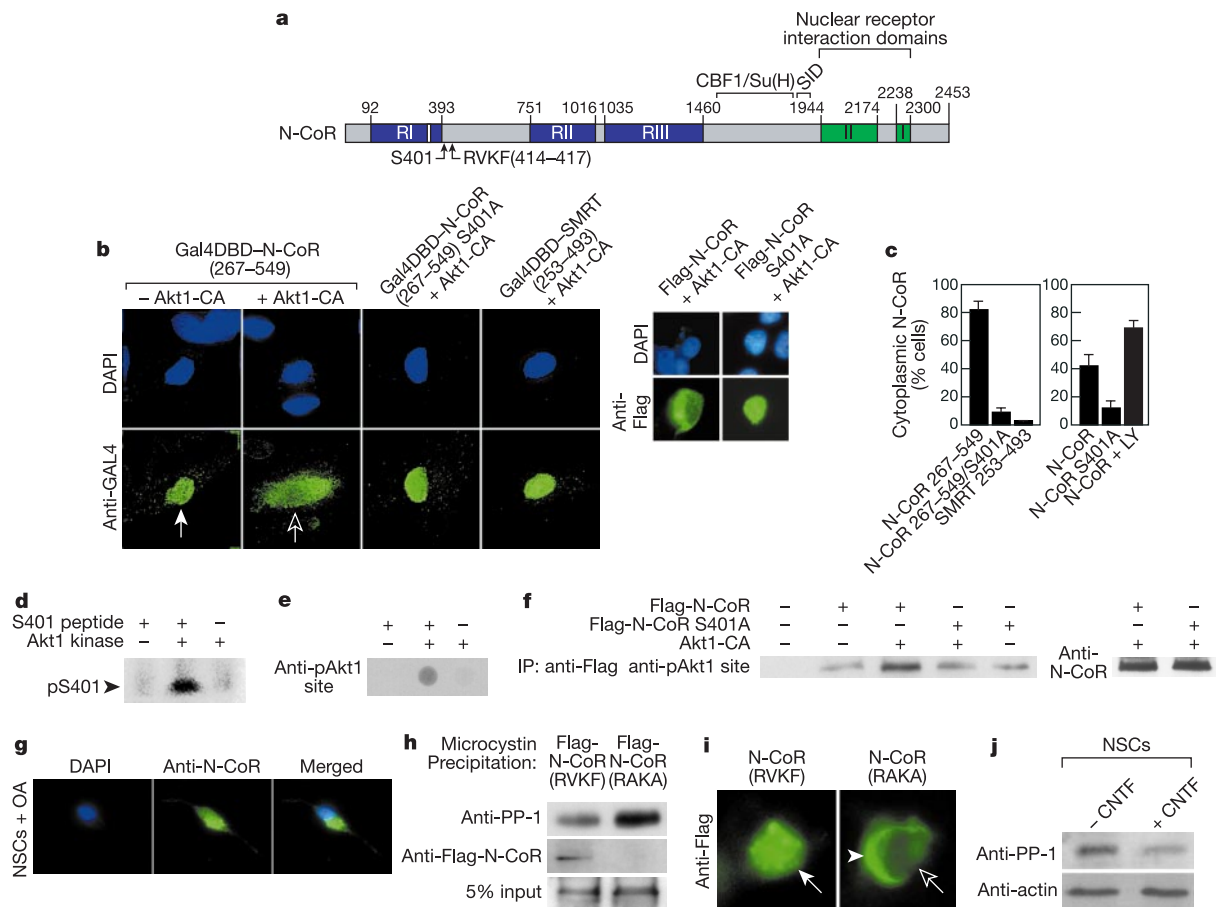


Figure 4 The subcellular localization of N-CoR is regulated through direct targeting by Akt1 kinase and protein phosphatase-1 (PP-1). **a**, Schematic model of N-CoR indicating putative regions for regulation by phosphorylation. **b**, Detection with anti-Gal4 or anti-Flag (green) antibody of transfected vectors as indicated in CV-1 cells in the absence or presence of constitutively active Akt1 kinase (Akt1-CA). Nuclear DAPI staining is indicated (blue). Closed arrow indicates nuclear N-CoR labelling. Open arrow indicates decrease in nuclear N-CoR labelling. **c**, Relative increase in cytoplasmic labelling of indicated constructs in CV-1 cells in the presence of Akt1-CA. **d**, *In vitro* phosphorylation of a synthetic 20-amino acid peptide harbouring the S401 site as assessed by incorporation of [γ -³²P]ATP (**d**), and an antibody against phosphorylated Akt1 kinase target sites (**e**). **f**, Immunoprecipitation of transfected HEK293 cells with anti-Flag antibody, followed by immunoblotting with an antibody raised against phosphorylated Akt1 kinase sites.

g, Endogenous N-CoR immunolabelling (green) in neural stem cells after 12 h of treatment with okadaic acid (OA) at concentrations inhibiting PP-1 activity (100 nM). Nuclear DAPI staining is indicated (blue). **h**, Whole-cell extracts from HEK293 cells transfected with indicated constructs were incubated with PP-1-interacting microcystin beads. After precipitation of the beads, immunoblotting of the co-precipitates was performed using anti-Flag or anti-PP-1 antibodies. **i**, Immunohistochemical detection of transfected constructs using an α -Flag antibody (green) in neural stem cells. Closed arrow indicates nuclear N-CoR labelling. Arrowhead indicates cytoplasmic N-CoR labelling. Open arrow indicates lack of nuclear N-CoR labelling. **j**, Immunoblotting of neural stem cell extracts with antibodies against PP-1 and actin in the absence of CNTF or 48 h after treatment with CNTF. PP-1 levels decreased to approximately 40% compared with control after CNTF treatment.

induced differentiation was always accompanied by translocalization and subsequent downregulation of endogenous N-CoR and induction of GFAP expression in approximately 25% of cells (Fig. 3e, and data not shown). Using a phosphorylation-specific antibody directed against Akt target sites, we detected prominent LY-sensitive phosphorylation of a protein with a relative molecular mass of 270,000 (M_r 270K) corresponding to endogenous N-CoR, 30 min after CNTF stimulation of neural stem cells (Fig. 3f). These experiments indicate a role for Akt1 kinase in neural differentiation, acting through regulation of N-CoR subcellular localization and protein levels in neural stem cells. Notably, brain-specific inactivation of the phosphatase PTEN, a negative regulator of Akt kinase, results in increased Akt1 activation accompanied by increased soma size and increased astrocytosis in certain parts of the brain (for review see ref. 21).

To determine whether Akt1 kinase directly regulates the nuclear export of N-CoR, ten overlapping regions of N-CoR spanning the entire complementary DNA linked to the Gal4 DNA-binding domain (Gal4DBD), which itself harbours an intact nuclear localization signal²², were co-transfected with Akt1-CA into CV-1 cells. Only one of these fusion proteins, Gal4DBD-N-CoR(267–549), showed significant increased cytoplasmic staining in the presence of Akt1-CA (83%; Fig. 4b, c). An algorithm²³ to reveal putative Akt1 kinase target sites used at low stringency predicted only one putative target site for Akt1 kinase in this region, centred at serine 401 (Fig. 4a). A synthetic 20-amino acid peptide harbouring the S401 site with flanking amino acids (ENNEKQMRQLSVIPPMFEDA) was efficiently phosphorylated by recombinant Akt1 kinase *in vitro* (Fig. 4d, e). A serine to alanine mutation at position 401 (S401A) of the Gal4DBD-N-CoR(267–549) fusion protein, as well as of the N-CoR holoprotein, significantly reduced the Akt1 kinase-mediated translocalization (Fig. 4b, c). We noted that the closely related co-repressor SMRT contains a conserved alanine residue at the corresponding site of serine 401 in N-CoR, and furthermore, the nuclear localization of a Gal4 fusion protein for the corresponding domain of SMRT (amino acids 253–493) was not altered in response to co-expression with Akt1-CA (Fig. 4b, c). Immunoprecipitation of Flag-tagged wild-type or S401A-mutated N-CoR from HEK293 cells using anti-Flag immunoglobulin- γ (IgG), followed by immunoblotting using the phospho-specific Akt1 kinase site antibody, revealed that wild-type N-CoR showed significantly increased phosphorylation (about threefold) in response to Akt1-CA, whereas N-CoR(S401A) showed no increased phosphorylation (Fig. 4f). Together, these data reveal a correlation between Akt1 activation, N-CoR phosphorylation, and the subcellular localization of N-CoR, although we cannot exclude the possibility that the effects of Akt1 on N-CoR could involve additional protein kinases acting as intermediaries.

Because the phosphorylation status of N-CoR correlated with its subcellular localization, we investigated the potential effects of regulated phosphatase activity. Notably, N-CoR contains a consensus interaction site for protein phosphatase-1 (PP-1), RVKF, at amino acids 414–417, which is just C-terminal to the Akt1 kinase site at S401 (Fig. 4a). This PP-1-binding site, as well as the Akt kinase target site and flanking residues, was found to be 100% conserved in N-CoR homologues from *Xenopus* to human. The activity and specificity of the catalytic subunit of PP-1 is determined by its recruitment to proteins harbouring this PP-1 interacting motif, RVXF^{24,25}. Treatment with PP-1-inhibiting concentrations (100 nM) of okadaic acid completely disrupted FGF2-dependent neural stem cell proliferation, and caused increased localization of N-CoR to the cytoplasm (Fig. 4g). Co-immunoprecipitation experiments using PP-1-interacting microcystin-coated beads revealed that PP-1 and N-CoR precipitated together *in vivo* (Fig. 4h). A construct harbouring a mutation of the PP-1 interaction site (RVKF $\rightarrow$ RAKA)²⁴ did not interact with PP-1 (Fig. 4h), exhibited aberrant, predominantly cytoplasmic subcellular localization (Fig. 4i), and increased basal phosphorylation (data not shown). Fur-

thermore, CNTF treatment downregulated PP-1 expression significantly in neural stem cells within 48 h (Fig. 4j). This suggests that CNTF coordinately downregulates PP-1 levels and activates Akt1 kinase, resulting in increased phosphorylation and cytoplasmic localization of N-CoR, thus decreasing its ability to maintain the neural stem cell state.

Although our data reveal a role for N-CoR-mediated repression in neural stem cells, it should be noted that N-CoR seems dispensable for early brain development and neural stem cell induction. The characteristics of neural progenitors show temporal variations, and in early development cytokine (CNTF/leukaemia inhibitory factor)-insensitive neuroepithelial cells constitute a major proliferating population in the forebrain^{1,26}. In these cells, alternative mechanisms of inhibition of glia differentiation are operating, including methylation of the *GFAP* promoter and high levels of neurogenic transcription factors that compete for the co-activators CBP/p300, which are required for glial differentiation (for review see ref. 27). However, in neural stem cells that have acquired cytokine sensitivity, *GFAP* gene methylation has ceased and the level of neurogenic transcription factors has decreased, and at this point FGF2-stimulated N-CoR repression becomes required for GFAP repression and the undifferentiated state. Thus, we suggest that N-CoR repression, under enzymatic regulation by Akt1 and PP-1, is a critical component of the FGF2-regulated genetic programme for maintaining a self-renewing, undifferentiated, multipotent neural stem cell state, and for repression of astroglia differentiation. $\square$

Methods

Neural stem cell cultures

Cortices from mice ($n = 120$) at embryonic day 12.5–13 were mechanically dispersed in Hanks balanced salt solution (Invitrogen), and plated individually at a density of 12,500 cells cm^{-2} to maintain a monolayer on 35-mm dishes coated with poly-L-ornithine and fibronectin (both Sigma). The primary cells were maintained in enriched N2 medium and generally received 15 ng ml^{-1} FGF2 (R&D Systems) every 8–10 h. CNTF or PDGF (R&D Systems) was administered at 20 ng ml^{-1} every 12 h. The culture of rat neural stem cells has been described in detail previously³ and was performed on 6–24-well and 35–150-mm plates in the presence of 10 ng ml^{-1} FGF2 added every 24 h. CNTF was administered at 10 ng ml^{-1} every 24 h.

Immunohistochemistry

Immunohistochemistry was performed largely as previously described¹¹. Cell cultures were fixed for 20 min with formalin. Antibody incubation was generally performed in a solution consisting of PBS with 1% BSA and 0.1% Triton X-100 at 4 °C overnight or for 1–2 h at room temperature. Apoptotic cell death was determined using Apoptag (Intergen) according to the supplier's recommendations.

Immunoblotting

Cells (neural stem cells, CV-1, HEK293) were rinsed in PBS, lysed in NETN buffer (150–300 mM NaCl) with protease inhibitor cocktail (Roche) and 1 mM phenylmethylsulphonyl fluoride (PMSF) for 15–20 min, and centrifuged for 15–20 min at 4 °C. After determination of protein concentration (Bio-Rad), cell lysates were used either for immunoprecipitation experiments (see below) or boiled for 5–10 min and loaded directly on to acrylamide (6–15%) gels in denaturing (SDS–polyacrylamide gel electrophoresis (PAGE)) conditions, followed by electrophoretic transfer (Bio-Rad). Western blotting was performed by standard protocols and developed using ECL reagents (Amersham) according to the supplier's recommendations.

Immunoprecipitation

Whole-cell lysates were incubated with antibody for 1–2 h at 4 °C on a rotating wheel followed by 1 h incubation with protein A/G agarose beads. For PP-1–N-CoR interaction studies, we used a microcystin-agarose 50% slurry (Upstate Biotechnology). Beads were subsequently centrifuged, washed, and the protein was eluted and subjected to immunoblotting.

Antibodies

Rabbit and mouse anti-GFAP (Biogenesis; Chemicon; ICN), rabbit anti-neomycin phosphotransferase II (Upstate Biotechnology), rabbit (gift from R. McKay) and mouse (Pharmingen) anti-nestin, mouse anti-A2B5 (a gift from A. Campagnoni), rabbit and guinea-pig anti-N-CoR²¹, rabbit anti-SMRT (ABR), rabbit and mouse anti-HA (Babco), mouse anti-Flag M2 (Sigma), rabbit anti-RBP-Jk/CBF1 (Chemicon), rabbit anti-Ki67, mouse anti-Gal4, rabbit anti-notch (Santa Cruz Biotechnology), rabbit anti-PP1 α and rabbit anti-phosphorylated Akt kinase target sites (Cell Signaling Technology) were used

with secondary antibodies from various sources (Molecular Probes, Chemicon, Vector Laboratories).

Pharmacological reagents

Treatment with pharmacological reagents occurred 30–60 min before growth factor or cytokine administration. The reagents were either present throughout the experiment or alternatively the medium was replaced after 6–12 h incubation. Reagents in this study were: LY294002 (10–50 μ M; Calbiochem) and okadaic acid (1 nM to 1 μ M; Calbiochem and Alomone Laboratories).

In vivo and in vitro kinase assays

The *in vivo* kinase assay was performed largely as recommended by the supplier (Akt Kinase Assay Kit, Cell Signaling Technology). *In vitro* kinase assays were performed using 1 $\times$ buffer (Akt Kinase Assay Kit, Cell Signaling Technology), 100 μ M ATP, 0.5 mM peptide (ENNEKQMRQLSVIPPMFMDA), 300 U ml⁻¹ Akt1 protein kinase. For radioactivity incorporation assay, 250 μ Ci ml⁻¹ [γ -³²P]ATP were included in the reaction. The reactions were quenched by addition of SDS gel buffer, and were either spotted on nitrocellulose membranes for immunoblotting, or were analysed by PAGE using 18% acrylamide gels.

Transfections and constructs

Transfections were performed with lipofection reagents (Effectene, Qiagen) largely following the supplier's recommendations. Typically, 1–2 μ g DNA was lipofected on a 60-mm dish for 6–12 h in the presence of FGF2, after which the cells received fresh medium with growth factors as indicated. Previously unpublished constructs used in this study are: pcDNA3-HA-N-CoR C-terminal (amino acids 1501–2300), pCMX-N-CoR(S401A)-Flag, pCMX-N-CoR-RAKA-Flag, pCMX-N-CoR Δ 1501–1800-Flag, and pCMX-Gal4-SMRT(253–493).

Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed on neural stem cells largely as described²⁸. Neural stem cells grown on 150-mm dishes were collected and crosslinked using 1% formalin for 10 min at room temperature, and the extracts were sonicated until the DNA fragments were 500–800 bp in size. Cell extracts were subsequently incubated with 5 μ g IgG or antibodies against N-CoR or CBP1 overnight at 4 °C. The extracts were incubated with protein A-sepharose beads (Sigma) for 1 h. After extensive washing of the beads, proteins were eluted and reversed by crosslinking for 6 h at 65 °C. After DNA purification, PCR was performed at 29–33 cycles. Primers were: GFAP, 5'-GACTAAGCTGTTTCTCCTCGGC-3' (sense), 5'-CAAGGTCACGTGATCCAGAG-3' (antisense); HES5, 5'-CGTGTCTCTTCTCCCATTTG-3' (sense), 5'-GATCCAGTGTGATCCGAGG-3' (antisense).

Received 13 June; accepted 23 September 2002; doi:10.1038/nature01156.
Published online 16 October 2002.

1. Temple, S. The development of neural stem cells. *Nature* **414**, 112–117 (2001).
2. Ghosh, A. & Greenberg, M. E. Distinct roles for bFGF and NT-3 in the regulation of cortical neurogenesis. *Neuron* **15**, 89–103 (1995).
3. Johe, K. K., Hazel, T. G., Muller, T., Dugich-Djordjevic, M. M. & McKay, R. D. Single factors direct the differentiation of stem cells from the fetal and adult central nervous system. *Genes Dev.* **10**, 3129–3140 (1996).
4. Palmer, T. D., Markakis, E. A., Willhoite, A. R., Safar, F. & Gage, F. H. Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS. *J. Neurosci.* **19**, 8487–8497 (1999).
5. Temple, S. & Alvarez-Buylla, A. Stem cells in the adult mammalian central nervous system. *Curr. Opin. Neurobiol.* **9**, 135–141 (1999).
6. Hörlein, A. J. *et al.* Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377**, 397–404 (1995).
7. Kao, H. Y. *et al.* A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. *Genes Dev.* **12**, 2269–2277 (1998).
8. Frisén, J. & Lendahl, U. Oh no, Notch again! *Bioessays* **23**, 3–7 (2001).
9. Faux, C. H., Turnley, A. M., Epa, R., Cappai, R. & Bartlett, P. F. Interactions between fibroblast growth factors and Notch regulate neuronal differentiation. *J. Neurosci.* **21**, 5587–5596 (2001).
10. Hitoshi, S. *et al.* Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. *Genes Dev.* **16**, 846–858 (2002).
11. Jepsen, K. *et al.* Combinatorial roles of the nuclear receptor corepressor in transcription and development. *Cell* **102**, 753–763 (2000).
12. Raballo, R. *et al.* Basic fibroblast growth factor (Fgf2) is necessary for cell proliferation and neurogenesis in the developing cerebral cortex. *J. Neurosci.* **20**, 5012–5023 (2000).
13. Rao, M. S., Noble, M. & Mayer-Proschel, M. A tripotential glial precursor cell is present in the developing spinal cord. *Proc. Natl Acad. Sci. USA* **95**, 3996–4001 (1998).
14. Choi, B. H. Prenatal gliogenesis in the developing cerebrum of the mouse. *Glia* **1**, 308–316 (1988).
15. Yasui, Y. *et al.* Roles of Rho-associated kinase in cytokinesis; mutations in Rho-associated kinase phosphorylation sites impair cytokinetic segregation of glial filaments. *J. Cell Biol.* **143**, 1249–1258 (1998).
16. Perissi, V. *et al.* Molecular determinants of nuclear receptor-corepressor interaction. *Genes Dev.* **13**, 3198–3208 (1999).
17. Besnard, F. *et al.* Multiple interacting sites regulate astrocyte-specific transcription of the human gene for glial fibrillary acidic protein. *J. Biol. Chem.* **266**, 18877–18883 (1991).
18. Kaneko, R., Hagiwara, N., Leader, K. & Sueoka, N. Glial-specific cAMP response of the glial fibrillary acidic protein gene cell lines. *Proc. Natl Acad. Sci. USA* **91**, 4529–4533 (1994).

19. Kahn, M. A. *et al.* Ciliary neurotrophic factor activates JAK/Stat signal transduction cascade and induces transcriptional expression of glial fibrillary acidic protein in glial cells. *J. Neurochem.* **68**, 1413–1423 (1997).
20. Bellacosa, A., Testa, J. R., Staal, S. P. & Tsichlis, P. N. A retroviral oncogene, akt, encoding a serine-threonine kinase containing an SH2-like region. *Science* **254**, 274–277 (1991).
21. Morrison, S. J. Pten-uating neural growth. *Nature Med.* **8**, 16–18 (2002).
22. Rhee, Y., Gurel, F., Gafni, Y., Dingwall, C. & Citovsky, V. A genetic system for detection of protein nuclear import and export. *Nature Biotechnol.* **18**, 433–437 (2000).
23. Yaffe, M. B. *et al.* A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nature Biotechnol.* **19**, 348–353 (2001).
24. Egloff, M. P. *et al.* Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1. *EMBO J.* **16**, 1876–1887 (1997).
25. Aggen, J. B., Nairn, A. C. & Chamberlin, R. Regulation of protein phosphatase-1. *Chem. Biol.* **7**, R13–R23 (2000).
26. Molne, M. *et al.* Early cortical precursors do not undergo LIF-mediated astrocytic differentiation. *J. Neurosci. Res.* **59**, 301–311 (2000).
27. Sauvageot, C. M. & Stiles, C. D. Molecular mechanisms controlling cortical gliogenesis. *Curr. Opin. Neurobiol.* **12**, 244–249 (2002).
28. Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D. & Broach, J. R. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.* **7**, 592–604 (1993).

Acknowledgements We are grateful to L. van Grunsven and R. McKay for their help in neural stem cell biology, and for providing advice throughout this study. Constructs and reagents were provided by D. Altomare, J. Testa, E. Lamar, C. Kintner, J. De Vellis, B. Andersen, T. Sugihara, D. Rose, and the Campagnoni Laboratory. We are also grateful to S. McMullen for microscopy assistance; C. Nelson and A. Krones for various cell culture assistance; H. Taylor for animal care; P. Myer for artwork; M. Fisher for assistance with the manuscript; V. Perissi for numerous reagents, discussions, and advice on ChIP; and members of the Rosenfeld laboratory for comments, discussions and various reagents, in particular L. Erkman, A. Gleiberman, V. Kumar, R. McEvilly and D. Solum. M.G.R. is an investigator with the Howard Hughes Medical Institute. O.H. was partially funded by the Swedish Brain Foundation. This work was funded by a grant from NIH.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.G.R. (e-mail: mrosenfeld@ucsd.edu).

Specific aspartyl and calpain proteases are required for neurodegeneration in *C. elegans*

Popi Syntichaki*, Keli Xu†‡, Monica Driscoll† & Nektarios Tavernarakis*

* Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology, Heraklion 71110, Crete, Greece

† Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, 604 Allison Road, Piscataway, New Jersey 08855, USA

Necrotic cell death underlies the pathology of numerous human neurodegenerative conditions¹. In the nematode *Caenorhabditis elegans*, gain-of-function mutations in specific ion channel genes such as the degenerin genes *deg-1* and *mec-4*, the acetylcholine receptor channel subunit gene *deg-3* and the G_s protein α -subunit gene *gsa-1* evoke an analogous pattern of degenerative (necrotic-like) cell death in neurons that express the mutant proteins^{2–6}. An increase in concentrations of cytoplasmic calcium in dying cells, elicited either by extracellular calcium influx or by release of endoplasmic reticulum stores, is thought to comprise a major death-signalling event^{7,8}. But the biochemical mechanisms by which calcium triggers cellular demise remain largely unknown. Here we report that neuronal degeneration inflicted by various genetic lesions in *C. elegans* requires the activity of the calcium-regulated CLP-1 and TRA-3 calpain proteases and aspartyl proteases ASP-3 and ASP-4. Our findings show that two

‡ Present address: Program in Developmental Biology, The Hospital for Sick Children, 555 University Avenue, Toronto Ontario, M5G 1X8, Canada.

distinct classes of proteases are involved in necrotic cell death and suggest that perturbation of intracellular concentrations of calcium may initiate neuronal degeneration by deregulating proteolysis. Similar proteases may mediate necrotic cell death in humans.

Neuronal degeneration initiated by hyperactive MEC-4, an ion channel subunit that is normally required for mechanotransduction in the six touch receptor neurons of *C. elegans*, is reminiscent of excitotoxic cell death in mammals^{9,10}. Electron microscopy studies of dying neurons in animals expressing a gain-of-function *mec-4* allele (*u231* or *d*; dominant) have shown extensive degradation of cellular contents during mid to late stages of cell death⁹. This ultrastructural feature suggests that proteolysis may be central to the biochemical mechanism underlying neuronal degeneration. The main executioner protease, caspase CED-3, which mediates programmed cell death (apoptosis) in *C. elegans*, and three additional CED-3-related proteases (CSP-1, CSP-2 and CSP-3) that are encoded in the nematode genome are not required for cell death induced by *mec-4(d)* (ref. 11; P.S. and N.T., unpublished data). This indicates that a distinct, non-apoptotic mechanism, which probably involves different proteases, functions in neurodegeneration in the nematode.

Aspartyl proteases are a class of catabolic hydrolases that, among others, includes lysosomal (cathepsin D) and non-lysosomal (cathepsin E) enzymes. We tested the requirement for aspartyl protease activity in neurodegeneration inflicted by hyperactive MEC-4 in touch receptor neurons in three ways. First, we used genetic backgrounds that have diminished levels of aspartyl protease activity. Nematode strains carrying mutations in three genes, *cad-1*, *daf-4* and *unc-52*, that encode otherwise unrelated proteins have been shown to maintain aspartyl protease activity that is 90% lower than in wild-type animals^{12–14}. We found that neurodegeneration induced by *mec-4(d)* was suppressed in *cad-1;mec-4(d)*, *daf-4;mec-4(d)* or *unc-52;mec-4(d)* mutant strains (Fig. 1a and Supplementary Information).

Second, we treated *mec-4(d)* animals with pepstatin A, an inhibitor of aspartyl proteases. Treatment ameliorated necrosis of the six touch receptor neurons (Fig. 1a). Third, we subjected *mec-4(d)* animals to starvation conditions under which aspartyl protease activity has been shown to drop to 5–10% that of well-fed nematodes^{12,15}. We observed a reduction in the number of dying neurons in the progeny of starved animals (Fig. 1a). Survival of touch receptor neurons was confirmed by the presence of cells expressing green fluorescent protein (GFP) from the *mec-4* promoter in adult animals.

Suppression of *mec-4(d)*-induced cell death in the genetic backgrounds and under the conditions examined was not a consequence of a reduction in the quantities of toxic MEC-4(d) protein. We used the *mec-4* promoter to drive expression of both *lacZ* and GFP reporter genes in *cad-1*, *daf-4* or *unc-52* mutant strains, as well as in animals treated with pepstatin A and starved animals. We did not detect any difference in reporter gene expression between these animals and well-fed wild-type animals (Fig. 1b, left, *cad-1(j1)* is shown as an example). Similarly, we assayed the relative stability of MEC-4 by using reporter genes with either *lacZ* or GFP fused at the carboxy terminus of the full-length MEC-4 protein. Quantities of protein were either not affected or even slightly increased by manipulations that reduced aspartyl protease activity (Fig. 1b, right, *cad-1(j1)* is shown).

We next determined whether aspartyl protease deficiency is generally protective against necrotic cell death. Gain-of-function (*d*) mutations in three otherwise unrelated genes, the degenerin *deg-1*, the α -7 nicotinic acetylcholine receptor Ca^{2+} channel subunit gene *deg-3*, and *gsa-1*, which encodes the $\text{G}\alpha_s$ subunit, trigger degeneration of specific sets of neurons expressing the toxic variants^{2,4–6}. We tested the general requirement for aspartyl protease in neurodegeneration by introducing various death-inducing

mutations into aspartyl-protease-deficient *cad-1*, *daf-4* and *unc-52* mutant strains. Cell death inflicted by toxic *deg-1(d)* and *deg-3(d)* alleles and by overexpressing the hyperactivated $\text{G}\alpha_s$ (Q227L) variant ($\alpha_s(\text{gf})$) was suppressed in genetic backgrounds deficient in aspartyl proteases. Starvation, which results in diminished aspartyl protease activity, also ameliorated neurodegeneration (Fig. 1c–e). Neuron survival was confirmed in adult animals by scoring expression of GFP. Similar to expression of *mec-4*, expression of *deg-1*, *deg-3* and $\alpha_s(\text{gf})$ was not reduced in protease-deficient genetic backgrounds or under starvation conditions. We conclude that aspartyl protease activity is required generally for neurodegeneration caused by deleterious mutations in

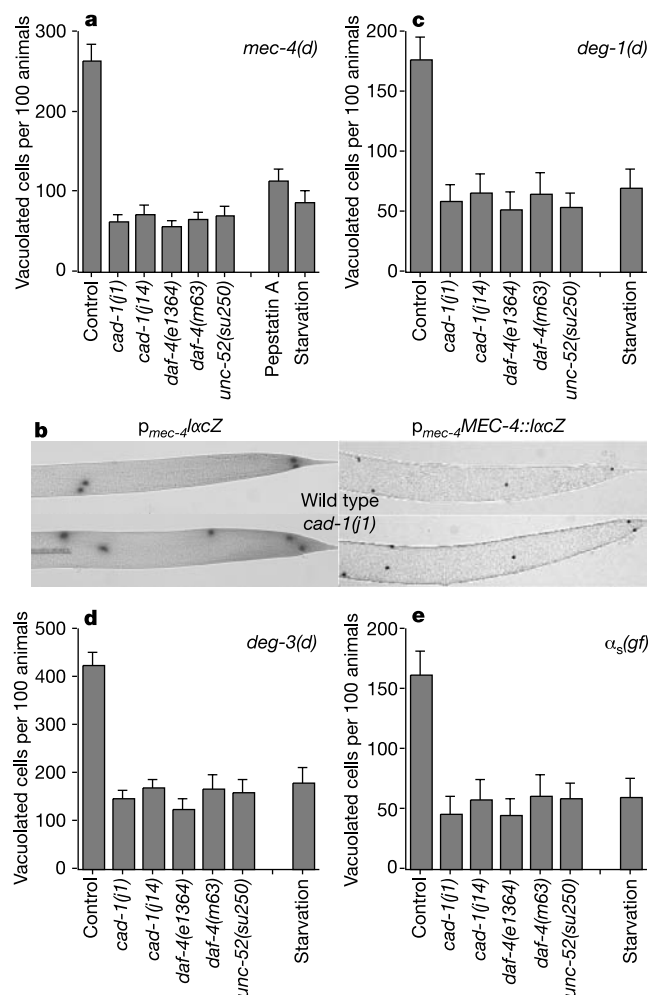


Figure 1 Aspartyl protease deficiency suppresses neurodegeneration in *C. elegans*. **a**, Number of vacuolated touch receptors at the L1 stage per 100 animals carrying the *mec-4(d)* allele in genetic backgrounds with reduced aspartyl protease activity, after treatment with pepstatin A and under conditions of starvation ($n = 250$, $P < 0.0001$, unpaired *t*-test). **b**, Expression of *lacZ* in touch receptor neurons driven solely by the *mec-4* promoter (left) or fused at the end of the full-length *mec-4* gene (right). Aspartyl protease deficiency in *cad-1(j1)* (bottom) does not affect *mec-4* expression or stability, as compared with the wild-type background (top). **c**, Vacuolated PVC interneurons at the L2 stage per 100 *deg-1(d)* animals in genetic backgrounds with reduced aspartyl protease activity and under conditions of starvation ($n = 250$, $P < 0.0001$, unpaired *t*-test). **d**, Vacuolated IL1 sensory neurons and PVC interneurons per 100 L1 *deg-3(d)* mutant larvae in genetic backgrounds with reduced aspartyl protease activity and under conditions of starvation ($n = 250$, $P < 0.001$, unpaired *t*-test). **e**, Vacuolated PVC interneurons at the L1 stage per 100 $\alpha_s(\text{gf})$ animals in genetic backgrounds with reduced aspartyl protease activity and under conditions of starvation ($n = 250$, $P < 0.0001$, unpaired *t*-test).

many different *C. elegans* genes.

There are at least six expressed aspartyl protease genes (*asp-1* to *asp-6*) encoded in the nematode genome (ref. 16; see Supplementary Information for multiple sequence alignment and phylogenetic tree). To identify those that contribute to the protease activity required for neurodegeneration, we systematically knocked down the expression of each *asp* gene by RNA interference (RNAi) in *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ genetic backgrounds. As a positive control in these experiments, we knocked down *crt-1*, which encodes calreticulin, an endoplasmic reticulum (ER) chaperone required for neurodegeneration induced by *mec-4(d)* and $\alpha_s(gf)$, but not *deg-3(d)* (ref. 8). Although RNAi is relatively ineffective for genes expressed in mature *C. elegans* neurons (ref. 17; N.T. and P.S.,

unpublished data), we observed suppression of neurodegeneration triggered by *mec-4(d)* in *crt-1(RNAi)* animals (Fig. 2a, c). It seems likely that RNAi is more effective with genes that are required at early developmental stages in the nervous system (degeneration occurs soon after the touch receptor neurons are born during late embryogenesis and the first larval stage in *mec-4(d)* *C. elegans* mutants⁹).

Of the six aspartyl protease genes examined, *asp-3* and *asp-4* were specifically required for neurodegeneration (Fig. 2a–c). *asp-1* knockdown also detectably reduced neurodegeneration but to a much lower extent. In a reciprocal approach, we introduced each *asp* gene into *cad-1(j1);mec-4(d)* double mutant animals, in which

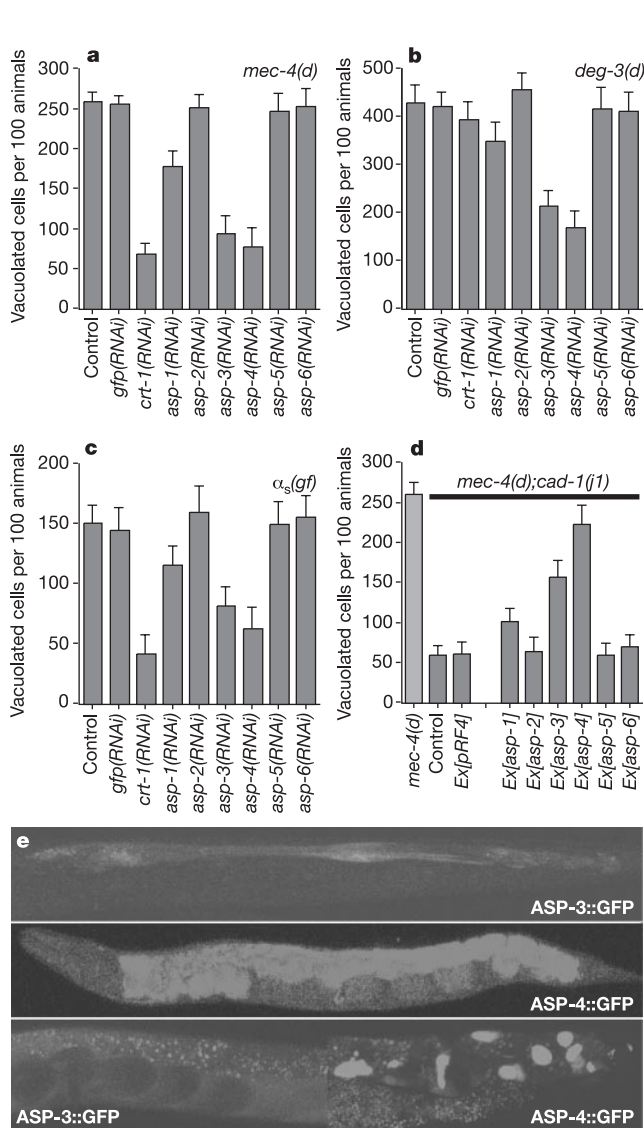


Figure 2 Specific aspartyl proteases are required for neurodegeneration in *C. elegans*. **a–c**, Effect of RNAi with the indicated *asp* genes in *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ mutants. RNAi with *crt-1* was used as a positive control; RNAi with *gfp* was used as a negative control. Interference with *asp-3* and *asp-4* results in significant suppression in all three genetic backgrounds ($n = 200$, $P < 0.0001$, unpaired *t*-test). Efficacy of RNAi was assessed as described in Methods. **d**, Degenerating touch receptors in transgenic *cad-1(j1);mec-4(d)* animals carrying each of the indicated *asp* genes on extrachromosomal arrays. *Ex[asp-3]* and *Ex[asp-4]* restore cell death in *cad-1(j1);mec-4(d)* double mutants ($n = 150$, $P < 0.0001$, unpaired *t*-test). **e**, Top and middle, images of animals expressing ASP-3::GFP and ASP-4::GFP. Bottom, confocal images of the subcellular localization of ASP-3::GFP and ASP-4::GFP (see Supplementary Information for details).

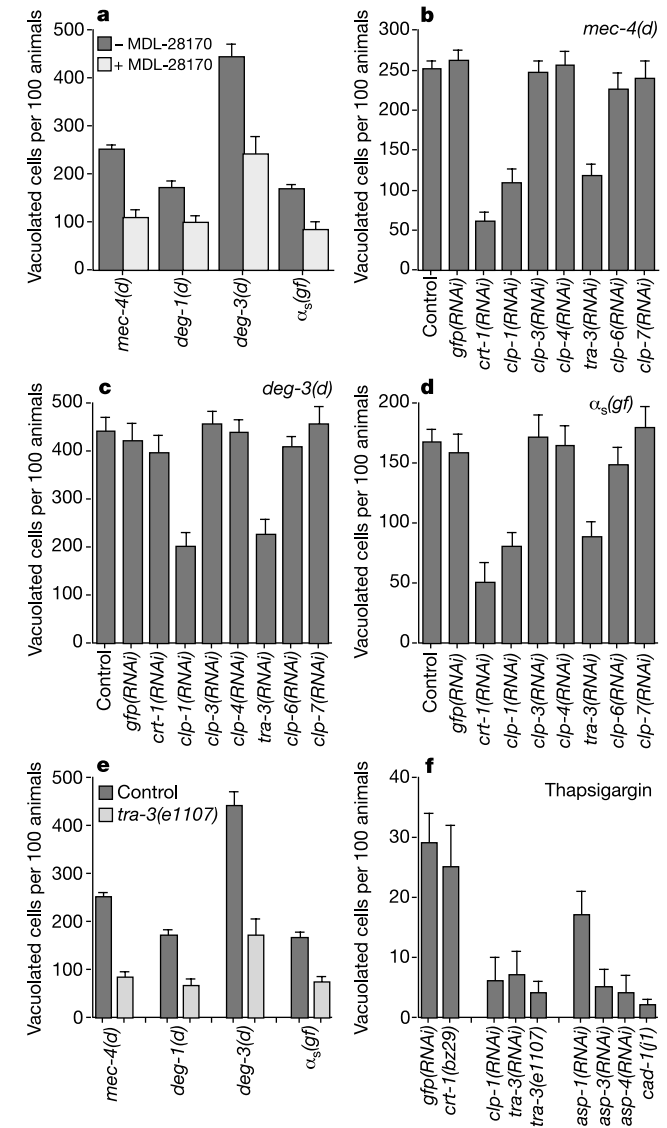


Figure 3 Specific calpain proteases are required for neurodegeneration in *C. elegans*. **a**, Degenerating neurons in *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ animals without or after treatment with calpain inhibitor MDL-28170 ($n = 150$, $P < 0.001$, unpaired *t*-test). **b–d**, Effect of RNAi with the indicated *clp* genes in *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ mutants. RNAi with *crt-1* was used as a positive control; RNAi with *gfp* was used as a negative control. Interference with *clp-1* and *clp-3* results in significant suppression in all three genetic backgrounds ($n = 200$, $P < 0.0001$, unpaired *t*-test). **e**, Neurodegeneration in single *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ mutants, and in double *tra-3(1107)* homozygotes originating from *tra-3(1107)* homozygous parents ($n = 100$, $P < 0.001$, unpaired *t*-test). **f**, Vacuolated cells per 100 L1 progeny of animals treated with thapsigargin and subjected to RNAi with the indicated genes ($n = 80$, $P < 0.001$, unpaired *t*-test).

neuronal degeneration caused by *mec-4(d)* is suppressed owing to the aspartyl protease deficiency of *cad-1(j1)*. We observed that degeneration was restored in animals carrying the *asp-3* and *asp-4* transgenes and to a much lower extent in animals carrying the *asp-1* transgene (Fig. 2d). Similarly, overexpression of *asp-3* and *asp-4* restored degeneration in *daf-4(e1364);mec-4(d)* and *unc-52(su250);mec-4(d)* double mutants (see Supplementary Information). Together, our results indicate that ASP-3 and ASP-4 aspartyl proteases are required for neurodegeneration inflicted by diverse genetic insults in *C. elegans*, but that ASP-1 contributes only marginally.

ASP-1 contains a conserved lysosome-targeting, *N*-glycosylation site (Asp 71) that is typical of cathepsin D lysosomal proteases that are predominantly localized to lysosomes¹⁶. Notably, ASP-3 and ASP-4 do not contain this *N*-glycosylation site but have another potential *N*-glycosylation site that is common in non-lysosomal

cathepsin E proteases (ref. 16 and see Supplementary Information). We examined the expression and subcellular localization of ASP-3 and ASP-4 by fusing GFP at the C termini of both proteins. Strong expression was observed in intestinal cells and to a much lesser extent in other types of cell, including muscle cells, the hypodermis and neurons (Fig. 2e). Both fusion proteins were localized mainly in the cytoplasm, but were also found in lysosomes that appear as distinct autofluorescent puncta (Fig. 2f).

Overexpression of caspase CED-3, the protease that mediates execution of programmed cell death, is sufficient to induce apoptosis in the absence of upstream death initiator signals¹⁸. We examined whether, by analogy, increased expression of the aspartyl proteases ASP-1, ASP-3 and ASP-4 is sufficient to inflict degeneration of specific neurons in *C. elegans*. We used the *mec-4* promoter to overexpress *asp-1*, *asp-3* and *asp-4* in the touch receptor neurons, and the motor neuron-specific *unc-8* promoter¹⁹ to drive overexpression in the ventral nerve cord motor neurons. A low percentage of transgenic animals expressing *asp-3* and *asp-4*, but not *asp-1*, in touch receptor neurons showed spontaneous vacuolation of these neurons during late embryogenesis and the early L1 larval stage and failed to respond normally to gentle touch as adults (for *asp-3*, $12.3 \pm 0.8\%$; for *asp-4*, $13.1 \pm 1.2\%$; $n = 250$, background is zero). Similarly, increased expression of *asp-3* and *asp-4* in motor neurons resulted in variably uncoordinated animals with vacuolated cells in the ventral nerve cord (for *asp-3*, $9.2 \pm 0.5\%$; for *asp-4*, $11.6 \pm 0.7\%$; $n = 250$, background is zero). Staining of cell nuclei with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI) in affected adults revealed missing neurons, indicating that overexpression of aspartyl protease caused neuronal death rather than mere malfunction (data not shown).

What is the mechanism that relays signals generated by upstream death initiating events to executioner aspartyl proteases? Perturbation of cytosolic calcium ($[Ca^{2+}]_i$) homeostasis has been implicated in necrotic cell death both in mammals and in *C. elegans*^{7,8}. But the mechanism by which Ca^{2+} contributes to cell death remains unclear. Calpains are diverse intracellular papain-like cysteine proteases that require Ca^{2+} for activation²⁰. In primate hippocampal neurons, degeneration after acute ischaemia is accompanied by an increase in $[Ca^{2+}]_i$ and concomitant activation of calpain. In addition, activated calpain seems to be localized to disrupted lysosomal membranes (reviewed in ref. 21). These findings have culminated in formulation of the 'calpain-cathepsin' hypothesis, whereby an increase in $[Ca^{2+}]_i$ activates calpains, which in turn mediate rupture of lysosomes and leakage of killer cathepsins that dismantle the cell^{1,21}.

To elucidate the role of calpain activity in *C. elegans* neurodegeneration, we treated *mec-4(d)*, *deg-1(d)*, *deg-3(d)* and $\alpha_s(gf)$ mutant animals with Z-Val-Phe-CHO (MDL-28170), a potent calpain inhibitor. Treatment markedly reduced the number of degenerating neurons in all four mutants without reducing the expression of *mec-4*, *deg-1*, *deg-3* or $\alpha_s(gf)$ (Fig. 3a). This observation suggested that calpain proteases are involved in the cell death process.

The *C. elegans* genome encodes 17 genes with similarity to calpain, 7 of which show significant identity to mammalian calpains over their whole length (*clp-1* to *clp-7*, see WormBase (<http://www.wormbase.org>) and Supplementary Information for multiple sequence alignment and phylogenetic tree). *clp-5* corresponds to *tra-3*, a previously characterized gene that is involved in *C. elegans* sex determination²². The TRA-3 protease is regulated by Ca^{2+} , but it lacks a typical calmodulin-like Ca^{2+} -binding domain²³. In TRA-3, a C2-like domain and two Ca^{2+} -binding sites in the protease core are probably Ca^{2+} sensors that activate this enzyme (see Supplementary Information). Examination of the other CLP sequences showed that CLP-1, CLP-2, CLP-6 and CLP-7 contain motifs that are typical of calpains, including a thiol (cysteine) protease active site and a Ca^{2+} -binding domain, whereas the other two lack either or both

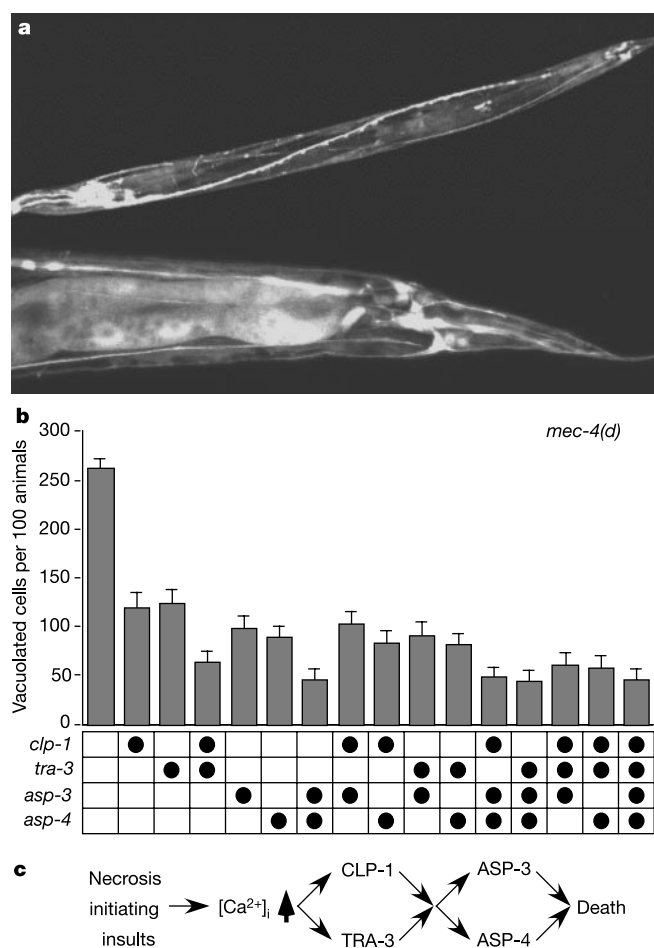


Figure 4 Calpains act sequentially with aspartyl proteases to facilitate cell death. **a**, Expression of *clp-1*. Top, image of animal expressing GFP from the *clp-1* promoter widely in the nervous system. Bottom, confocal image detailing expression in the intestine and neurons of the tail ganglion (see also Supplementary Information). **b**, Neurodegeneration assays in *mec-4(d)* animals subjected to RNAi with protease genes. *In vitro* synthesized dsRNAs, corresponding to calpain genes *clp-1* and *tra-3* and aspartyl proteases *asp-3* and *asp-4*, were injected either individually or in pools as indicated. Bars represent the average of three independent experiments ($n = 300$). Knockdown of both calpain or both aspartyl protease genes results in significantly more extended quenching of neurodegeneration than for any single gene ($P < 0.001$, unpaired *t*-test). **c**, Working model for a necrotic cell death pathway in *C. elegans*. Death triggering signals are relayed to a specific set of executioner aspartyl proteases through sensing of $[Ca^{2+}]_i$ perturbations by calpains.

(see Supplementary Information for multiple sequence alignment). We investigated the role of calpain proteases in neurodegeneration by RNAi-mediated knockdown of *clp-1*, *clp-3*, *clp-4*, *clp-6*, *clp-7* and *tra-3* in *mec-4(d)*, *deg-3(d)* and $\alpha_s(gf)$ mutant strains. Cell death was suppressed in all three strains when animals were subjected to RNAi with *clp-1* and *tra-3* but not *clp-3*, *clp-4*, *clp-6* or *clp-7* (Fig. 3b–d). We obtained similar results with *tra-3;mec-4(d)*, *tra-3;deg-1(d)*, *tra-3;deg-3(d)* and *tra-3; $\alpha_s(gf)$* double mutants (Fig. 3e), further confirming the requirement for TRA-3 in neurodegeneration. Expression of *mec-4*, *deg-3* and $\alpha_s(gf)$ genes was not detectably reduced in these experiments.

Studies have suggested that release of Ca^{2+} from ER stores to the cytoplasm contributes to neurodegeneration initiated by hyperactive MEC-4 or $G\alpha_s$ (ref. 8). Forced release of Ca^{2+} from ER stores by treatment with thapsigargin, a compound that also inhibits the ER Ca^{2+} re-uptake pump SERCA and results in a net increase in $[Ca^{2+}]_i$, induces necrotic cell death in *C. elegans*⁸. We found that calpain activity is required for thapsigargin-induced cell death: treatment with thapsigargin was not effective in animals subjected to RNAi with *clp-1*, or in *tra-3* mutants (Fig. 3f). We also examined the requirement for aspartyl protease activity in thapsigargin-induced cell death. *cad-1* mutants and animals subjected to RNAi with *asp-3* or *asp-4* were resistant to the toxic effects of thapsigargin, but RNAi with *asp-1* slightly ameliorated cell death. Loss of calreticulin function in *crt-1(bz29)* null mutants, which blocks neurodegeneration induced by *mec-4(d)* and $\alpha_s(gf)$, but not *deg-3(d)*, did not suppress thapsigargin toxicity (Fig. 3f). These observations indicate that although CRT-1 is required for the build-up of noxious $[Ca^{2+}]_i$, calpain and aspartyl proteases function downstream of $[Ca^{2+}]_i$ signalling to facilitate death.

clp-1 is expressed strongly in many types of cell and tissue, including muscle cells and neurons (Fig. 4a and Supplementary Information). *tra-3* is also expressed in the nervous system of the animal (S. Sokol and P. Kuwabara, personal communication). We examined synergy between proteases of the same type, as well as between aspartyl proteases and calpains, in neurodegeneration. We observed that simultaneous RNAi with both *asp-3* and *asp-4* resulted in an enhanced suppression of neurodegeneration induced by *mec-4(d)*. Similarly, RNAi with both calpains further increased suppression. But we did not observe synergy between aspartyl proteases and calpains (Fig. 4b). Therefore, aspartyl and calpain proteases function in the same pathway that facilitates neurodegeneration in *C. elegans*. We did not achieve complete or near-complete blockage of neurodegeneration in these experiments. Incomplete suppression of cell death by aspartyl or calpain protease deficiency in our trials could be due to the limited capacity of RNAi to knockdown genes efficiently in the nematode nervous system, the contribution of other additional biochemical pathways and protease activities, or both. A comprehensive study of the remaining calpain proteases including CLP-2, which contains all of the catalytic residues that are typical of calpains (Supplementary Table 1), may illuminate this issue.

We propose that diverse death-initiating conditions converge, in part, to increase $[Ca^{2+}]_i$, which signals the activation of calpain proteases that subsequently engage executioner lysosomal and cytoplasmic aspartyl proteases, leading to cell destruction (Fig. 4c). Consistent with this model, neurodegeneration inflicted by cell-specific overexpression of *asp-3* and *asp-4* cannot be bypassed by a deficiency in either or both calpains (Supplementary Table 2). The identification of two specific classes of protease that are required for neurodegeneration in *C. elegans* may provide insight into similar pathologies in mammals. The lysosomal degradation system has been found to be upregulated in neurons of individuals affected with Alzheimer's disease²⁴, and cathepsin D expression is induced under conditions of excitotoxic cell death (reviewed in ref. 25). In addition, calpain inhibitors can be protective in certain cases of nerve or muscle degeneration after ischaemic

episodes^{26,27}. These findings suggest that, similar to apoptosis, necrotic cell death mechanisms are conserved from nematodes to humans, and they highlight specific executioner proteases as potential targets for therapeutic intervention in an effort to battle neurodegenerative disorders. □

Methods

Strains and genetics

We used standard procedures for *C. elegans* maintenance, crosses and other genetic manipulations²⁸. The cultivation temperature was kept at 20 °C, unless noted otherwise. We used the following strains: wild-type N2 Bristol isolate, *cad-1(j1)II*, *cad-1(j14)II*, *csf-1(nr2018)II*, *unc-52(su250)II*, *daf-4(e1364)III*, *daf-4(m63)III*, *tra-3(e1107)/dpy-4(e1166)IV*, *deg-3(u662)V*, referred to in the text as *deg-3(d)*; *deg-1(u38)X*, referred to in the text as *deg-1(d)*; *mec-4(u231)X*, referred to in the text as *mec-4(d)*; and *nuls5[p_{glt-1}Gα_s(Q227L) p_{glt-1}GFP]*, referred to in the text as $\alpha_s(gf)$ (ref. 6). The following double mutants were examined for neurodegeneration: *cad-1(j1)II;mec-4(u231)X*, *cad-1(j14)II;mec-4(u231)X*, *cad-1(j1)II;deg-3(u662)V*, *cad-1(j1)II; $\alpha_s(gf)$* , *csf-1(nr2018)II;mec-4(u231)X*, *csf-1(nr2018)II;deg-3(u662)V*, *csf-1(nr2018)II; $\alpha_s(gf)$* , *unc-52(su250)II;mec-4(u231)X*, *unc-52(su250)II;deg-1(u38)X*, *unc-52(su250)II;deg-3(u662)V*, *unc-52(su250)II; $\alpha_s(gf)$* , *daf-4(e1364)III;mec-4(u231)X*, *daf-4(m63)III;mec-4(u231)X*, *daf-4(e1364)III;deg-1(u38)X*, *daf-4(e1364)III;deg-3(u662)V*, *daf-4(e1364)III; $\alpha_s(gf)$* , *tra-3(e1107)IV;mec-4(u231)X*, *tra-3(e1107)IV;deg-3(u662)V*, *tra-3(e1107)IV; $\alpha_s(gf)$* , *tra-3(e1107)/+IV;mec-4(u231)X*, *tra-3(e1107)/+IV;deg-3(u662)V* and *tra-3(e1107)/+IV; $\alpha_s(gf)$* .

The *tra-3(e1107)/+* heterozygotes segregate *tra-3(e1107)* homozygotes, which, owing to maternal effects are phenotypically wild type and segregate abnormal sterile males or intersex²². We assayed cell death in *tra-3(e1107)* homozygotes originating from *tra-3(e1107)/+* heterozygotes and in abnormal progeny of *tra-3(e1107)* homozygotes, as well as in *tra-3(e1107)/+* heterozygotes. Neurodegeneration was suppressed only in the abnormal progeny of *tra-3(e1107)* homozygotes, indicating the maternal contribution of TRA-3 in cell death. *uvEx(asp-nrol-6(su1006));cad-1(j1)II;mec-4(u231)* (referred to in the figure legend as *Ex[asp-n] cad-1(j1)II;mec-4(u231)*, where *n* is 1–6) were constructed by injecting PCR fragments encompassing each *asp* gene, together with plasmid pRF4 carrying the dominant transformation marker *rol-6(su1006)*, into the gonads of *cad-1(j1)II;mec-4(u231)* gravid hermaphrodites.

Plasmid constructs and RNA interference

We generated reporter constructs by fusing GFP at the C terminus of ASP-3 and ASP-4, and at the second exon of *clp-1*. The *mec-4* and *unc-8* promoters were fused upstream of *asp-1*, *asp-3* and *asp-4* coding sequences to achieve ectopic overexpression of aspartyl protease genes in touch receptor neurons and motor neurons, respectively. Overexpression of *clp-1* and *tra-3* using a similar strategy did not result in detectable neurodegeneration (data not shown). For RNAi experiments, we generated plasmid constructs for *in vitro* synthesis of double-stranded RNAs (dsRNAs) corresponding to various aspartyl and calpain proteases that were subsequently injected as described²⁹. There is no sequence identity in the DNA between the protease genes tested, which eliminates incident dsRNA cross-interference in our experiments.

We carried out neurodegeneration assays (see below) in the L1 progeny of injected individuals. In a complementary approach, we also constructed plasmids for synthesis of dsRNAs in *Escherichia coli* bacteria that were fed to animals as described³⁰. In all cases, injecting dsRNA resulted in the strongest interference effects in the nervous system. We assayed the effectiveness of RNAi with aspartyl protease gene expression by measuring aspartyl protease activity in total animal extracts¹² and by monitoring expression of full-length *asp::GFP* reporter genes. RNAi with *tra-3* phenocopied the *tra-3* loss-of-function phenotype, confirming effective interference.

Cell death assays

Animals were staged using the criteria of developmental timing and extent of gonadal development. Neurodegeneration was generally scored during the mid-L1 larval stage. We prepared L1 larvae by washing gravid adult populations of plates and allowing the remaining eggs to hatch for 4–5 h. For the strains expressing *mec-4(d)*, and ASP-1, ASP-3 and ASP-4 from the *mec-4* promoter, we scored neurodegeneration by the characteristic vacuolated appearance of the six touch receptor neurons using differential interference contrast (Nomarski) microscopy. For *deg-1(d)* and $\alpha_s(gf)$, we scored vacuolation of the two PVC interneurons in the tail, and for *deg-3(d)*, we scored vacuolation of IL1 sensory neurons in the head and PVC interneurons in the tail. For strains expressing ASP-1, ASP-3 and ASP-4 from the *unc-8* promoter, we assayed vacuolation of ventral nerve cord motor neurons. Cell death or survival was confirmed by the lack or presence of GFP expression in these neurons and DAPI staining of cell nuclei, in adult animals. Thapsigargin-induced cell death was assayed as described⁸. Animals were treated with 10 μ M ml⁻¹ thapsigargin. We injected protease inhibitors (5 μ M pepstatin A, Sigma; 10 μ M MDL-28170, Calbiochem) in the body cavity of gravid adults and assayed neurodegeneration in the progeny of injected individuals. We imposed starvation by moving well-fed young adults to plates devoid of bacteria and allowing them to lay eggs that were hatched in the absence of food. Neurodegeneration was assayed in the L1 progeny of starved animals. For *unc-52(su250)*, *daf-4(e1364)* and *daf-4(m63)* alleles, which are temperature-sensitive, we assayed neurodegeneration at 25 °C, where effects were maximal. Statistical analysis was carried out using the Microsoft OfficeXP Excel software package.

Received 3 July; accepted 21 August 2002; doi:10.1038/nature01108.

1. Syntichaki, P. & Tavernarakis, N. Death by necrosis: uncontrollable catastrophe, or is there order behind the chaos? *EMBO Rep.* **3**, 604–609 (2002).
2. Chalfie, M. & Wolinsky, E. The identification and suppression of inherited neurodegeneration in *Caenorhabditis elegans*. *Nature* **345**, 410–416 (1990).
3. Driscoll, M. & Chalfie, M. The *mec-4* gene is a member of a family of *Caenorhabditis elegans* genes that can mutate to induce neuronal degeneration. *Nature* **349**, 588–593 (1991).
4. Treinin, M. & Chalfie, M. A mutated acetylcholine receptor subunit causes neuronal degeneration in *C. elegans*. *Neuron* **14**, 871–877 (1995).
5. Korswagen, H. C., Park, J. H., Ohshima, Y. & Plasterk, R. H. An activating mutation in a *Caenorhabditis elegans* G_s protein induces neural degeneration. *Genes Dev.* **11**, 1493–1503 (1997).
6. Berger, A. J., Hart, A. C. & Kaplan, J. M. G_α-induced neurodegeneration in *Caenorhabditis elegans*. *J. Neurosci.* **18**, 2871–2880 (1998).
7. Sattler, R. & Tymianski, M. Molecular mechanisms of calcium-dependent excitotoxicity. *J. Mol. Med.* **78**, 3–13 (2000).
8. Xu, K., Tavernarakis, N. & Driscoll, M. Necrotic cell death in *C. elegans* requires the function of calreticulin and regulators of Ca²⁺ release from the endoplasmic reticulum. *Neuron* **31**, 957–971 (2001).
9. Hall, D. H. *et al.* Neuropathology of degenerative cell death in *Caenorhabditis elegans*. *J. Neurosci.* **17**, 1033–1045 (1997).
10. Lee, J. M., Zipfel, G. J. & Choi, D. W. The changing landscape of ischaemic brain injury mechanisms. *Nature* **399**, A7–A14 (1999).
11. Chung, S., Gumienny, T. L., Hengartner, M. O. & Driscoll, M. A common set of engulfment genes mediates removal of both apoptotic and necrotic cell corpses in *C. elegans*. *Nature Cell Biol.* **2**, 931–937 (2000).
12. Jacobson, L. A. *et al.* Identification of a putative structural gene for cathepsin D in *Caenorhabditis elegans*. *Genetics* **119**, 355–363 (1988).
13. Rogalski, T. M., Mullen, G. P., Bush, J. A., Gilchrist, E. J. & Moerman, D. G. UNC-52/perlecan isoform diversity and function in *Caenorhabditis elegans*. *Biochem. Soc. Trans.* **29**, 171–176 (2001).
14. Estevez, M. *et al.* The *daf-4* gene encodes a bone morphogenetic protein receptor controlling *C. elegans* dauer larva development. *Nature* **365**, 644–649 (1993).
15. Hawdon, J. M., Emmons, S. W. & Jacobson, L. A. Regulation of proteinase levels in the nematode *Caenorhabditis elegans*. Preferential depression by acute or chronic starvation. *Biochem. J.* **264**, 161–165 (1989).
16. Tcherepanova, I., Bhattacharyya, L., Rubin, C. S. & Freedman, J. H. Aspartic proteases from the nematode *Caenorhabditis elegans*. Structural organization and developmental and cell-specific expression of *asp-1*. *J. Biol. Chem.* **275**, 26359–26369 (2000).
17. Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A. & Driscoll, M. Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nature Genet.* **24**, 180–183 (2000).
18. Miura, M., Zhu, H., Rotello, R., Hartwig, E. A. & Yuan, J. Induction of apoptosis in fibroblasts by IL-1 beta-converting enzyme, a mammalian homolog of the *C. elegans* cell death gene *ced-3*. *Cell* **75**, 653–660 (1993).
19. Tavernarakis, N., Shreffler, W., Wang, S. & Driscoll, M. *unc-8*, a DEG/ENAC family member, encodes a subunit of a candidate mechanically gated channel that modulates *C. elegans* locomotion. *Neuron* **18**, 107–119 (1997).
20. Sorimachi, H. & Suzuki, K. The structure of calpain. *J. Biochem. (Tokyo)* **129**, 653–664 (2001).
21. Yamashima, T. Implication of cysteine proteases calpain, cathepsin and caspase in ischemic neuronal death of primates. *Prog. Neurobiol.* **62**, 273–295 (2000).
22. Barnes, T. M. & Hodgkin, J. The *tra-3* sex determination gene of *Caenorhabditis elegans* encodes a member of the calpain regulatory protease family. *EMBO J.* **15**, 4477–4484 (1996).
23. Sokol, S. B. & Kuwabara, P. E. Proteolysis in *Caenorhabditis elegans* sex determination: cleavage of TRA-2A by TRA-3. *Genes Dev.* **14**, 901–906 (2000).
24. Adamec, E., Mohan, P. S., Cataldo, A. M., Vonsattel, J. P. & Nixon, R. A. Up-regulation of the lysosomal system in experimental models of neuronal injury: implications for Alzheimer's disease. *Neuroscience* **100**, 663–675 (2000).
25. Ferri, K. F. & Kroemer, G. Organelle-specific initiation of cell death pathways. *Nature Cell Biol.* **3**, E255–E263 (2001).
26. Stracher, A. Calpain inhibitors as therapeutic agents in nerve and muscle degeneration. *Ann. NY Acad. Sci.* **884**, 52–59 (1999).
27. Huang, Y. & Wang, K. K. The calpain family and human disease. *Trends Mol. Med.* **7**, 355–362 (2001).
28. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
29. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
30. Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* **2**, research0002.1–0002.10 (2001).

Supplementary Information accompanies the paper on Nature's website
(<http://www.nature.com/nature>).

Acknowledgements We thank P. Kuwabara and S. Sokol for communicating results before publication; C. Samara for technical help; and colleagues for discussions and comments on the manuscript. Some nematode strains used in this work were provided by the Caenorhabditis Genetics Center, which is funded by the NIH National Center for Research Resources. Certain strains were from the NemaPharm Group of Axys Pharmaceuticals. We thank A. Fire for plasmid vectors and L. Jacobson for *cad-1* mutant strains. This work was funded in part by grants from NIH NINDS to M.D. and by grants from HFSP and IMBB to N.T.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.T.
(e-mail: tavernarakis@imbb.forth.gr).

Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase iota

Ahmad Faii*, Said Aoufouchi*, Eric Flatter*, Quentin Guéranger, Claude-Agnès Reynaud & Jean-Claude Weill

INSERM U373, Faculté de Médecine Necker-Enfants Malades, 156 rue de Vaugirard, 75730, Paris Cedex 15, France

* These authors contributed equally to this work

Somatic hypermutation of immunoglobulin genes is a unique, targeted, adaptive process. While B cells are engaged in germinal centres in T-dependent responses, single base substitutions are introduced in the expressed *VH/VL* genes to allow the selection of mutants with a higher affinity for the immunizing antigen. Almost every possible DNA transaction has been proposed to explain this process, but each of these models includes an error-prone DNA synthesis step that introduces the mutations^{1,2}. The Y family of DNA polymerases³—pol η, pol ι, pol κ and rev1—are specialized for copying DNA lesions and have high rates of error when copying a normal DNA template^{4,5}. By performing gene inactivation in a Burkitt's lymphoma cell line inducible for hypermutation, we show here that somatic hypermutation is dependent on DNA polymerase iota.

Burkitt's lymphoma cell lines represent *in vitro* models of the hypermutation of immunoglobulin (Ig) genes. Mutation either can occur constitutively, as in the Ramos cell line⁶, or can be induced, as in the BL2 cell line, after the engagement of several surface receptors—a stimulus that resembles the process *in vivo*^{7,8}. In the BL2 cell line, mutations are induced in the G1 phase of the cell cycle, occur on one DNA strand of its rearranged *VH* gene, and eventually become fixed by replication in one of the daughter cells⁹. The feasibility of gene targeting in the BL2 cell line has been shown by inactivation of the *activation-induced cytidine deaminase* (*AID*) gene, which totally abolishes hypermutation in BL2 (ref. 9), as it does *in vivo*^{10,11}.

We used the same gene inactivation procedure to generate BL2 clones deficient in polymerase iota (pol ι), one of the two human homologues of yeast RAD30, which shows short-gap filling, highly error-prone polymerase activity^{12,13,14}. Figure 1a shows the organization of the human *POL*1 exons and the constructs used to inactivate both alleles. Two pol-ι-null clones (54 and 267) were generated from a pol-ι heterozygous clone. The two clones had the same proliferation rate as the original BL2 cell line, with a similar mitotic index and cell-cycle profile, and showed no chromosomal aberrations or translocations in spreads of metaphase cells (data not shown). No pol ι expression was detected when cell extracts from the 54 and 267 ι-null clones were immunoblotted using rabbit polyclonal antibodies raised against human pol ι (Fig. 1b). Expression of pol ι was restored by transfecting the cells with pIRES vectors driving the expression of the complete human pol-ι complementary DNA under the control of the promoter for cytomegalovirus (CMV; Fig. 1b): within the variations inherent to quantification by western blot, expression was comparable to that observed in the normal BL2 cell line. In spite of several attempts, we could not obtain clones overexpressing pol ι, either in a normal or in a pol-ι-null background. We selected two restored clones for each targeted cell for further analysis.

The BL2 cell line can be induced for hypermutation of its rearranged *VH* gene by two different procedures. In the first, crosslinking IgM in co-culture with a T-helper (T_H) clone or T_H cell line results in the occurrence of mutations after 3 d in culture^{7,15}. We have adapted this somatic hypermutation (SHM) induction

system by using the CB15 cell line—a T_H cell transformed by the herpes virus Saimiri—in the co-culture, and IgM crosslinking followed by aggregation of the biotinylated antibody against IgM by streptavidin beads. In the second procedure, the crosslinking of three surface receptors (IgM, CD19, CD21), followed by the same aggregation of the biotinylated anti-IgM antibody by streptavidin beads results in the occurrence of mutations observed after 90 min of incubation⁹. The mutation frequency resulting from a single round of mutagenesis is similar in both procedures ($4\text{--}6 \times 10^{-4}$ mutations per base pair; Table 1). The BL2 cell line has also a very low constitutive mutation frequency of its rearranged *VH* gene when maintained in culture; this frequency does not exceed 10^{-4} mutations per base pair after 2–3 months in culture.

No detectable increase in mutation frequency could be observed when the 54 or 267 pol- ι -null clones were induced for hypermutation by either the T-cell co-culture or the three-antibodies assay (Table 1). There were seven mutations in 298 *V* sequences after stimulation, compared with five mutations out of 222 sequences before stimulation. By contrast, the induction of mutations was restored on re-expression of the pol- ι cDNA in these clones to frequencies comparable to those observed in the normal BL2 cell line (Fig. 1b). There were 83 mutations in 547 *V* sequences after stimulation, compared with 6 mutations in 547 sequences before stimulation. In addition, the pattern of mutation restored in these pol- ι -expressing clones was similar to the pattern in wild-type BL2, with a high targeting of GC base pairs and a bias for transitions (Table 2 and Supplementary Fig. 1). As for the normal BL2 cell line, the *C μ* gene of the pol- ι -restored clones was not targeted for mutation after induction (one mutation in 71 *C μ* gene sequences

Table 1 **V gene mutations in pol- ι -null and pol- ι -restored BL2 subclones**

Cell type	Unstimulated	Stimulated	
		Anti-IgM/CD19/CD21	Anti-IgM/CB15
BL2	0.71×10^{-4} (7/236)	3.7×10^{-4} (29/187)	3.8×10^{-4} (26/165)
pol- ι -null clones			
54	0.82×10^{-4} (3/88)	0.96×10^{-4} (4/100)	0.51×10^{-4} (2/95)
267	0.36×10^{-4} (2/134)	1.00×10^{-4} (1/24)	$<0.30 \times 10^{-4}$ (0/79)
pol- ι -null clones restored with pol- ι cDNA			
54-7	0.31×10^{-4} (1/78)	2.9×10^{-4} (14/95)	3.1×10^{-4} (9/69)
54-10	0.43×10^{-4} (2/91)	2.4×10^{-4} (8/79)	ND
267-20	0.27×10^{-4} (1/89)	4.0×10^{-4} (19/93)	5.1×10^{-4} (22/104)
267-12	0.50×10^{-4} (2/97)	4.0×10^{-4} (9/67)	ND

Mutation frequencies are expressed as mutations per nucleotide, with the total number of mutations per number of *V4-39-JH5* genes (416 bp) sequenced given in parentheses.

of 399 base pairs (bp); data not shown).

Expression of *AID* varies among Ramos subclones kept in culture and is correlated with their variable mutation frequencies¹⁶. We therefore monitored the concentration of AID protein in the pol- ι -null clones. AID expression was similar to that in the restored clones, thus excluding the possibility that our results might be due to fluctuations in AID expression (Fig. 1b). Although a 3-d T-cell co-culture induction assay has shown that pol- ι messenger RNA increases specifically in BL2 cells after 12 h of culture¹⁵, we did not detect any increase in expression of pol- ι protein in the short time assay. The same result was observed with expression of AID protein, indicating that induction of SHM does not require the enhanced synthesis of either of these proteins (Fig. 2).

Pol ι is a distributive enzyme that works efficiently on short DNA gaps and can generate most of the mutations observed in SHM when it copies normal DNA. It behaves as an A/T mutator in the middle of the template and as a G/C mutator at the end when acting on a primer terminus with a long template overhang, and can extend all types of mispair by a one-nucleotide addition¹³. It also has a deoxyribose phosphate lyase activity¹⁷ that might be involved in an early step of the process, after an initial entry in DNA mediated by a glycosylase and/or endonuclease attack.

In contrast to inactivation of the *AID* gene, which abolishes both the constitutive and the induced mutation frequency in BL2 cells⁹, inactivation of pol ι did not seem to affect the background mutation frequency, which raises the issue of whether other polymerases might contribute to this low constitutive mutagenic activity without being mobilized in the induction process. It has been proposed that AID may act as a cytidine deaminase on genomic DNA and thus generate G•U mispairs¹⁸. Such mismatches, if not repaired, would generate a G/C to A/T transition after semiconservative replication. In the 12 mutations collected from 520 *V* sequences of pol- ι -null clones, eight out of ten base substitutions were G/C to A/T transitions (Supplementary Fig. 1; the other two were frameshift mutations in the (G)₈ tract in CDR3). Thus, these background

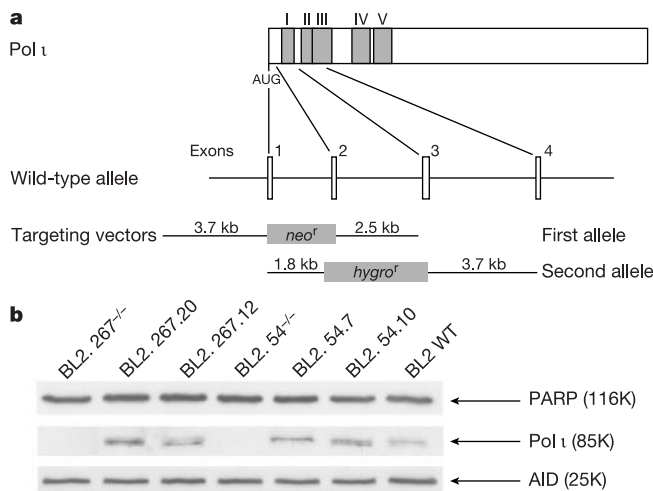


Figure 1 Inactivation of the gene encoding pol ι in the BL2 cell line and restoration of protein expression by transfection with pol- ι cDNA. **a**, Gene targeting strategy for pol ι . The five conserved domains described for all members of the pol Y family are represented in the pol- ι protein sequence^{5,14}; the corresponding position of each genomic exon is indicated (the straight line represents the beginning of the first four exons, except exon 1, for which it marks the position of the ATG initiation codon). For the first allele, *neo^r* replaces exons 1 and 2, which contain the ATG initiator codon and the domain I (DLD motif), respectively¹⁴. For the second allele, *hygro^r* replaces exon 2 and interrupts exon 3. **b**, Analysis of pol- ι protein expression in pol- ι -null and pol- ι -restored BL2 subclones. Pol ι and expression of AID was monitored in the following BL2 clones: two pol- ι -null clones (54, 267), four pol- ι -restored clones obtained by transfection of the two pol- ι -null clones with pol- ι -expressing vectors (54-7, 54-10, 267-12, 267-20) and wild-type BL2 as a control. Cell lysates were analysed by immunoblotting and probed with antibodies against human pol ι ; the upper part and lower part of the membrane were subsequently incubated with antibodies against poly(ADP-ribose) polymerase (PARP) and against human AID, respectively. PARP was included as a loading control. The relative molecular masses of the proteins are indicated in parentheses.

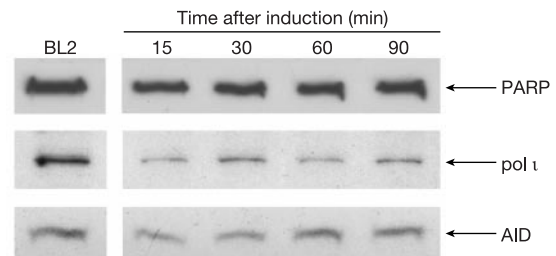


Figure 2 No increase in pol- ι and AID protein is observed during induction of SHM. BL2 cells were induced to undergo SHM by the three-antibodies crosslinking procedure. Cells were collected at the indicated time points, and protein extracts were analysed by immunoblotting with antibodies against pol- ι and AID as described in Fig. 1. PARP was included as a loading control.

Table 2 **V gene base substitution pattern in pol- ι -restored BL2 subclones**

From:	To:	A	T	G	C	Total
A			3.6 (0.9)	7.2 (4.6)	1.8 (1.4)	24.9 (20.5)
T		4.6 (3.1)		0 (2.7)	7.7 (7.8)	
G		28.1 (26.3)	13.5 (4.8)		1.3 (7.5)	75.1 (79.5)
C		9.6 (10.6)	19.0 (24.2)	3.6 (6.1)		
Transitions:		62.0 (62.9)		Transversions:	38.0 (37.1)	

Seventy-four base substitutions targeted in 416 bp of the VDJ sequence were analysed from a total of 83 mutations including nine insertions (listed in Table 1). The percentage of nucleotide substitution is corrected for the base composition of the V segment. Values are for the mutations induced in pol- ι -restored clones; for comparison, values for the induction of mutations in normal BL2 cells are given in parentheses (taken from ref. 9).

mutations could be essentially generated by AID.

It is clear that the kinetics of mutagenesis that B cells experience in germinal centres (10^{-3} mutations per base pair per cell generation) is much closer to that of our induction assay than to the background mutation rate of BL2. Even the Ramos cell line, which has a much higher constitutive mutation rate, requires about 6 weeks of culture to accumulate a number of mutations in its V sequence comparable to that observed during induction in the BL2 cell line¹⁹. It has been proposed that pol η contributes to SHM from the observation of a reduced targeting to A/T base pairs in XP-V individuals, who are deficient in pol η ²⁰, and from the specific spectrum of mutations it generates *in vitro*^{21,22}. The BL2 cell line has a pattern of mutation that is biased towards G/C base pairs (80% of total mutations, Table 2) as compared with the normal SHM pattern *in vivo* (55% G/C in human VH genes²³), suggesting that there might be reduced expression of pol η in this cell. When assayed by polymerase chain reaction with reverse transcription (RT-PCR), no specific defect in pol- η expression could be detected in BL2 during the short induction time, as compared with the expression of other DNA polymerases involved in translesion synthesis (TLS) (data not shown). It is difficult at this stage to understand the cause of the G/C bias in BL2—a bias that is also observed in other lymphoma cell lines that are constitutive or inducible for SHM^{6–8,24}—and whether it might be due to a defective participation of pol η in the mutasome. Because pol η and pol ι arose by duplication of an ancestral RAD30 gene from which they are equally divergent³, they may partially complement each other under specific conditions. The other DNA polymerase proposed to be involved in SHM is pol ζ ^{25,26}, a TLS polymerase that is not particularly error-prone²⁷ but might be required to enable DNA synthesis, either at a nick that is a poor substrate for pol ι ¹³ or beyond a mismatch owing to a misincorporated nucleotide²⁷.

Our study shows that pol ι is a component of the immunoglobulin mutasome. It remains highly probable, as suggested previously^{20–22,25,26}, that other TLS polymerases may also participate in the process and may compensate for each other in particular *in vivo* situations. If this is true, unravelling the whole picture may require the cumulative inactivation of all of these enzymes. □

Methods

Induction of V gene hypermutation in the BL2 cell line

V gene hypermutation was induced by two different procedures. The first involved crosslinking of the surface IgM, CD19 and CD21 receptors, as described⁹. The second involved co-culture with the CB15 cell line, a T_H cell line transformed by the herpes virus Saimiri²⁸. We cultured CB15 cells in RPMI 1640 medium (Invitrogen) supplemented with 10% fetal calf serum (Hyclone), penicillin/streptomycin (Invitrogen) and 50 units ml⁻¹ interleukin-2 (Sigma). The cells were irradiated at 4,000 rad, resuspended in complete medium plus 2% Cytosol (CTS, Lyon), and plated at 500,000 cells per well in 24-well plates that had been coated overnight with antibodies against CD3 (2.5 μ g per well; OKT3, Janssen-Cilag). BL2 cells (2×10^6) were incubated in 0.5 ml of RPMI medium for 20 min in ice with a biotinylated antibody against IgM (0.6 mg ml⁻¹; Caltag), and then crosslinked with streptavidin-conjugated magnetic beads (Dynabeads M280, Dynal) for 20 min on a rotating wheel in the cold. We resuspended BL2 cells at 100,000 cells ml⁻¹ in complete RPMI medium supplemented with 2% Cytosol, and plated 0.5 ml into each well of a 24-well plate containing irradiated CB15 and antibodies against CD3 (at a BL2:CB15 ratio of 1:10). Cells were collected after 3 d of incubation at 37 °C, DNA was extracted using the

DNeasy tissue kit (Qiagen), and mutations were analysed after amplification of the V4-39/JH5 gene as described⁹. We identified mutations through multiple alignments of electropherograms using the AutoAssembler software (Applied).

Inactivation of the POLI gene in BL2

The human pol- ι genomic sequence was retrieved from the 180-kilobase (kb) HTGS clone AC021325, which contains all of the coding exons. DNA fragments used to generate the targeting constructs were amplified from BL2 genomic DNA and cloned using the TOPO-XL PCR cloning kit (Invitrogen), excised using *Sall* and *XhoI* if these sites had been included in the PCR primers or as *MluI*–*NotI* fragments if not, and cloned into the corresponding sites of a pBSK vector (Stratagene) to which *SfiI* and *MluI* sites had been added in the *SacI* site of its polylinker. Neomycin (*neo*^r) or hygromycin (*hygro*^r) resistance genes were inserted in between 5' and 3' flanking fragments.

We used the following PCR primers for amplification. First allele: 5' fragment, 5'– ι 1–5' (5'–CAGTCGAGTACCAGCCGCTCTGGTGTGGT–3') and 5'– ι 1–3' (5'–CAGTCGAGTACAGTCTCCAGTCGCTC–3', 3.1 kb); 3'–fragment, 3'– ι 1–5' (5'–CAGTCGAGTCAAATC CAGAGCTAAAG–3') and 3'– ι 1–3' (5'–CAGTCGAGTCAATGTCAGGTAACCAAC–3', 2.5 kb). Second allele: 5'–fragment, 5'– ι 2–5' (5'–CAGTCGAGAGAGCGACTGGAGACTG TAG–3') and 5'– ι 2–3' (5'–ACGTCGACAGGAAGAATGCAGAGTACG–3', 1.8 kb); 3'–fragment, 3'– ι 2–5' (5'–CTCAGTCGACGGTGGTTACCTGCAACTATG–3') and 3'– ι 2–3' (5'–CCAAGTCTCTCAACAACCTGG–3', 3.8 kb); the *Sall* and *XhoI* restriction sites are underlined. Pfu turbo polymerase (Stratagene) was used for amplifying the 5' fragment of the first construction and Herculase (Stratagene) for the 3' fragment of the second construction (30 s at 94 °C, 30 s at 62 °C, and 12 min at 68 °C for 40 cycles), the Expand Long Template PCR system (Roche) for the 3' fragment of the first construction (40 cycles according to the supplier's conditions which include an increment in the elongation time), and the Advantage 2 PCR kit (Clontech) for the 5' fragment of the second construction (30 s at 94 °C, 30 s at 64 °C and 4 min at 68 °C). Transfection of 30 μ g of *NotI*- or *MluI*-linearized constructs was done as described⁹.

We screened for gene targeted clones by PCR on pools of ten clones with the following primers located outside the construction and within the *neo*^r or *hygro*^r resistance gene: first allele, ι –5'–test, 5'–ATGACCCAAGCTCAAGACAGC–3' and *neo*^r, 5'–CATAGCGTTGGCTACCCGCTG–3' in 5', and 3'– ι 2–3' and *neo*^r reverse, 5'–CACGGGTAGCCACGCTATG–3' in 3'; second allele, ι –5'–test and *hygro*^r, 5'–GCTGTGTAGAAGTACTCGCC–3' (Expand Long Template PCR system, Roche). For the first allele, homologous recombination was obtained in one clone among 518 transfectants. For the second allele, 2 clones out of 535 were properly targeted. We attribute this low frequency to the use of low-fidelity PCR enzymes in the construction of the targeting vectors: sequence differences between this construct and the targeted gene are likely to reduce the rate of homologous recombination. Such an effect of strain-specific polymorphisms on homologous recombination was observed in mouse embryonic stem cells, in which a 0.6% sequence difference results in a 50-fold decrease in targeting efficiency²⁹. Accordingly, other gene inactivations carried out in BL2 with constructions amplified only with Pfu turbo DNA polymerase (Stratagene), an enzyme whose error frequency is below 10^{-4} per base pairs in the amplification conditions used, gave a targeting frequency in the 1–2% range (ref. 9 and our own unpublished data).

Transfection of pol- ι -null BL2 clones

We transfected the pol- ι -null clones with pol- ι -expressing vectors, which we constructed by amplification of a full-length human pol- ι cDNA (ref. 14; a gift from R. Woodgate) with the following primers, 5'– ι ACGACGAGGAAGACG, and 3'– ι , TACGCTTTGTGCCAG, and cloning in the *EcoRV* site of the pIRES puro vector (Invitrogen). Transfection and selection of clones were done as described⁹.

Analysis of BL2 clones by RT-PCR and western blot

Western blots were done as described⁹. Polyclonal antibodies against pol- ι (raised against a carboxy-terminal 15-mer KLH-conjugated peptide) were a gift from R. Woodgate.

Received 4 July; accepted 20 August 2002; doi:10.1038/nature01117.

- Gearhart, P. J. & Wood, R. D. Emerging links between hypermutation of antibody genes and DNA polymerases. *Nature Rev. Immunol.* **1**, 187–192 (2001).
- Weill, J.-C. *et al.* Ig gene hypermutation: a mechanism is due. *Adv. Immunol.* **80**, 183–202 (2002).
- Ohmori, H. *et al.* The Y-family of DNA polymerases. *Mol. Cell* **8**, 7–8 (2001).
- Friedberg, E. C., Wagner, R. & Radman, M. Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science* **296**, 1627–1630 (2002).
- Goodman, M. F. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **71**, 17–50 (2002).
- Sale, J. E. & Neuberger, M. S. TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. *Immunity* **9**, 859–869 (1998).
- Denepoux, S. *et al.* Induction of somatic mutation in a human B cell line in vitro. *Immunity* **6**, 35–46 (1997).
- Zan, H. *et al.* Induction of Ig somatic hypermutation and class switching in a human monoclonal IgM⁺IgD⁺ B cell line in vitro: definition of the requirements and modalities of hypermutation. *J. Immunol.* **162**, 3437–3447 (1999).
- Failli, A. *et al.* AID-dependent somatic hypermutation occurs as a DNA single strand event in the BL2 cell line. *Nature Immunol.* **3**, 815–821 (2002).
- Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
- Revy, P. *et al.* Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565–575 (2000).
- Tissier, A., McDonald, J. P., Frank, E. G. & Woodgate, R. Pol ι , a remarkably error-prone human DNA polymerase. *Genes Dev.* **14**, 1642–1650 (2000).
- Frank, E. G. *et al.* Altered nucleotide misinsertion fidelity associated with pol ι -dependent

- replication at the end of a DNA template. *EMBO J.* **20**, 2914–2922 (2001).
14. McDonald, J. P. *et al.* Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase η . *Genomics* **60**, 20–30 (1999).
 15. Poltoratsky, V. *et al.* Expression of error-prone polymerases in BL2 cells activated for Ig somatic hypermutation. *Proc. Natl Acad. Sci. USA* **98**, 7976–7981 (2001).
 16. Zhang, W. *et al.* Clonal instability of V region hypermutation in the Ramos Burkitt's lymphoma cell line. *Int. Immunol.* **13**, 975–984 (2001).
 17. Bebenek, K. *et al.* 5'-Deoxyribose phosphate lyase activity of human DNA polymerase ι in vitro. *Science* **291**, 2156–2159 (2001).
 18. Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99–103 (2002).
 19. Papavasiliou, F. N. & Schatz, D. G. The activation-induced deaminase functions in a postcleavage step of the somatic hypermutation process. *J. Exp. Med.* **195**, 993–998 (2002).
 20. Zeng, X. *et al.* DNA polymerase ϵ is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nature Immunol.* **2**, 537–541 (2001).
 21. Rogozin, I. B., Pavlov, Y. I., Bebenek, K., Matsuda, T. & Kunkel, T. A. Somatic mutation hotspots correlate with DNA polymerase ϵ error spectrum. *Nature Immunol.* **2**, 530–536 (2001).
 22. Pavlov, Y. I. *et al.* Correlation of somatic hypermutation specificity and A-T base pair substitution errors by DNA polymerase η during copying of a mouse immunoglobulin κ light chain transgene. *Proc. Natl Acad. Sci. USA* **99**, 9954–9959 (2002).
 23. Dörner, T., Foster, S. J., Brezinschek, H.-P. & Lipsky, P. E. Analysis of the targeting of the hypermutational machinery and the impact of subsequent selection on the distribution of nucleotides changes in human V(D)J rearrangements. *Immunol. Rev.* **162**, 161–171 (1998).
 24. Bachl, J. & Wabl, M. An immunoglobulin mutator that targets G:C base pairs. *Proc. Natl Acad. Sci. USA* **93**, 851–855 (1996).
 25. Zan, H. *et al.* The translesion DNA polymerase ζ plays a major role in Ig and bcl-6 somatic hypermutation. *Immunity* **14**, 643–653 (2001).
 26. Diaz, M., Verkoczy, L. K., Flajnik, M. F. & Klinman, N. R. Decreased frequency of somatic hypermutation and impaired affinity maturation but intact germinal center formation in mice expressing antisense RNA to DNA polymerase ζ . *J. Immunol.* **167**, 327–335 (2001).
 27. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S. & Prakash, L. Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature* **406**, 1015–1019 (2000).
 28. Mitrucker, H. W., Muller-Fleckenstein, I., Fleckenstein, B. & Fleischer, B. Herpes virus saimiri-transformed human T lymphocytes: normal functional phenotype and preserved T cell receptor signalling. *Int. Immunol.* **5**, 985–990 (1993).
 29. de Wind, N., Dekker, M., Berns, A., Radman, M. & te Riele, H. Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* **82**, 321–330 (1995).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank B. Fleckenstein for the CB15 cell line; H. Mossafa for the cytogenetic analysis of pol- ι -null clones; R. Woodgate for pol- ι -specific reagents; M. Radman for critically editing the manuscript. This work was supported by the Fondation Princesse Grace de Monaco and the Ligue Nationale Française contre le Cancer (Equipe labellisée). A.F. and S.A. have been supported during part of this work by the Fondation de France (Fondation contre la Leucémie). C.-A.R. and J.-C.W. share senior authorship.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C.W. (e-mail: weill@necker.fr).

Dual regulation of voltage-gated calcium channels by PtdIns(4,5)P₂

Li Wu*, Claudia S. Bauer*, Xiao-guang Zhen, Cheng Xie & Jian Yang

Department of Biological Sciences, Columbia University, New York, New York 10027, USA

* These authors contributed equally to this work

Voltage-gated calcium channels (VGCCs) conduct calcium into cells after membrane depolarization and are vital for diverse biological events¹. They are regulated by various signalling pathways, which has profound functional consequences^{1,2}. The activity of VGCCs decreases with time in whole-cell and inside-out patch-clamp recordings³. This rundown reflects persistent intrinsic modulation of VGCCs in intact cells. Although several mechanisms have been reported to contribute to rundown of L-

type channels^{3–6}, the mechanism of rundown of other types of VGCC is poorly understood. Here we show that phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂), an essential regulator of ion channels and transporters^{7–14}, is crucial for maintaining the activity of P/Q- and N-type channels. Activation of membrane receptors that stimulate hydrolysis of PtdIns(4,5)P₂ causes channel inhibition in oocytes and neurons. PtdIns(4,5)P₂ also inhibits P/Q-type channels by altering the voltage dependence of channel activation and making the channels more difficult to open. This inhibition is alleviated by phosphorylation by protein kinase A. The dual actions of PtdIns(4,5)P₂ and the crosstalk between PtdIns(4,5)P₂ and protein kinase A set up a dynamic mechanism through which the activity of VGCCs can be finely tuned by various neurotransmitters, hormones and trophic factors.

P/Q-type Ca²⁺ channels were expressed in *Xenopus* oocytes and macroscopic currents were recorded in giant membrane patches with Ba²⁺ as the external permeant ion. Because depolarization to +100 mV activated nearly all channels (see below), the tail current after such a depolarization (Fig. 1a, *I*_{tail(+100)}) could be used as a faithful readout of the total number of functional channels and its change with time as a true measure of the time course of rundown. Channel activity stayed constant in the cell-attached mode but waned on excision of inside-out patches (Fig. 1b). The time for *I*_{tail(+100)} to decay to 50% of its initial value (*T*₅₀ of rundown) was 5.9 ± 3.2 min (*n* = 25).

To test the effect of PtdIns(4,5)P₂ on rundown, we manipulated the membrane concentration of PtdIns(4,5)P₂ by either applying a brief pulse of PtdIns(4,5)P₂ or sequestering it with a specific antibody (refs 10, 15). A 2-min application of PtdIns(4,5)P₂ immediately after patch excision greatly slowed rundown (Fig. 1b), increasing the *T*₅₀ by 5.4-fold to 32 ± 9.9 min (*n* = 6; Fig. 1c). PtdIns(4,5)P₂ also increased *I*_{tail(+100)} when applied after some channels had undergone rundown (Fig. 1d), suggesting that it can reactivate, at least partially, channels that have been rundown already. Conversely, reducing the amount of PtdIns(4,5)P₂ with an antibody against PtdIns(4,5)P₂ accelerated rundown (Fig. 1e). An antibody specific for PtdIns(4)P had no effect at the same concentration. These results suggest that PtdIns(4,5)P₂ stabilizes the activity of P/Q-type channels and its depletion induces rundown.

Notably, P/Q-type channels underwent rundown more slowly when 2 mM Mg-ATP was added to the intracellular solution (Fig. 1c). But this stabilizing effect was not affected by the inclusion of either PKI 5-24, a potent inhibitor of protein kinase A (PKA), or the catalytic subunit of PKA and okadaic acid (an inhibitor of phosphatases 1 and 2A; Fig. 1c), suggesting that other signalling pathways than PKA are involved. A possibility is that Mg-ATP activates membrane-associated lipid kinases to replenish PtdIns(4,5)P₂ (ref. 8). Regardless of the underlying mechanism, sequestering PtdIns(4,5)P₂ with antibodies against PtdIns(4,5)P₂ still markedly accelerated rundown, and this could be partially reversed by subsequent application of PtdIns(4,5)P₂ (Fig. 1f). Together, these results indicate that PtdIns(4,5)P₂ is indispensable for maintaining P/Q-type channel activity in inside-out patches.

Unexpectedly, PtdIns(4,5)P₂ also exerted a voltage-dependent inhibitory effect (Fig. 2a): it inhibited the current elicited by weak depolarizations strongly and the current by intermediate depolarizations moderately, but it had little effect on the outward K⁺ current (flowing through Ca²⁺ channels) evoked by large depolarizations. The activation and peak current voltages were both shifted in the positive direction by 15–20 mV in the presence of PtdIns(4,5)P₂, but the reversal potential was unaffected (Supplementary Information).

The voltage dependence of PtdIns(4,5)P₂ inhibition could be seen more clearly in the tail currents after various depolarizations. In control solution, the tail current after depolarization to +10 mV (*I*_{tail(+10)}) or +100 mV (*I*_{tail(+100)}) decayed in parallel after patch

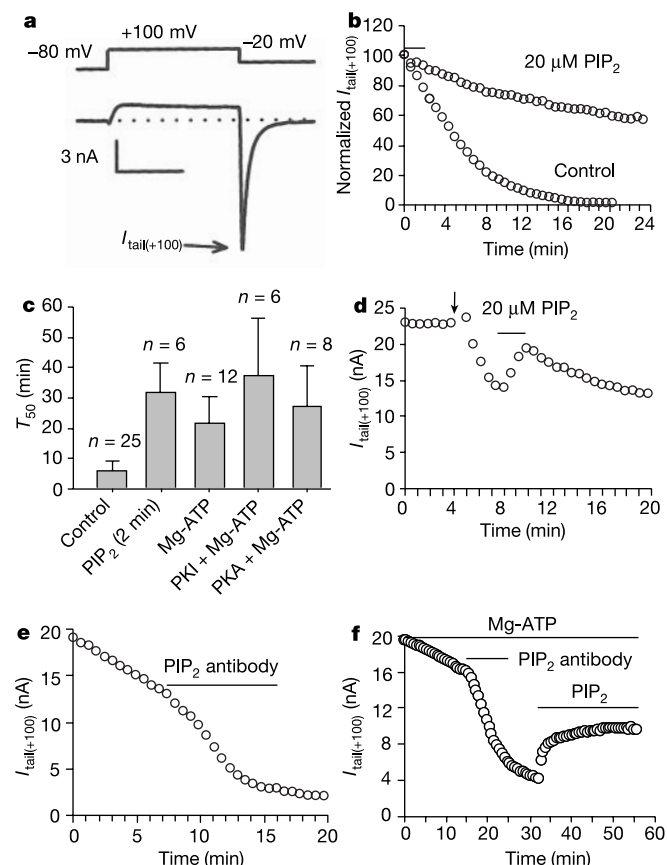


Figure 1 Stabilization of P/Q-type Ca^{2+} channel activity by PtdIns(4,5) P_2 (PIP_2). **a**, Representative current trace evoked by the indicated voltage protocol. **b**, Time course of $I_{\text{tail}(+100)}$ in control solution or with a 2-min PtdIns(4,5) P_2 treatment (unbroken line). Current is normalized to that obtained immediately after patch excision. **c**, T_{50} for decay of $I_{\text{tail}(+100)}$ in control solution, with a 2-min PtdIns(4,5) P_2 treatment (20 μM) immediately after patch excision, in 2 mM Mg-ATP, 2 mM Mg-ATP plus 5 μM PKI 5-24, or 2 mM Mg-ATP plus 8.8 units ml^{-1} PKA catalytic subunit plus 0.3 μM okadaic acid. **d**, Time course of $I_{\text{tail}(+100)}$ with a 2-min PtdIns(4,5) P_2 treatment during rundown. Arrow indicates patch excitation. **e**, **f**, Acceleration of $I_{\text{tail}(+100)}$ rundown by antibodies specific for PtdIns(4,5) P_2 (28.5 $\mu\text{g ml}^{-1}$ after 1:10 dilution) in control solution (**e**) or in the presence of 2 mM Mg-ATP, with partial reactivation of rundown channels by PtdIns(4,5) P_2 (**f**). Current sign is reversed in this and following figures.

excision (Fig. 2b) and the activation curve measured 8–10 min thereafter showed only a slight shift (Fig. 2c). A 2-min application of PtdIns(4,5) P_2 caused the tail currents to bifurcate (Fig. 2d): $I_{\text{tail}(+10)}$ was inhibited immediately and rapidly to $\sim 25\%$ of its starting value, whereas $I_{\text{tail}(+100)}$ was not obviously affected. (The residual $I_{\text{tail}(+10)}$ became very steady after stopping the application of PtdIns(4,5) P_2 , supporting the notion that PtdIns(4,5) P_2 retards rundown.) This result suggests that PtdIns(4,5) P_2 profoundly alters the voltage dependence of channel activation, making channels harder to open by weak depolarizations. The activation curve measured during a 10-min application of PtdIns(4,5) P_2 showed a gradual positive shift, which reached a maximum 8–10 min during application of PtdIns(4,5) P_2 when $I_{\text{tail}(+10)}$ was inhibited completely (Fig. 2e).

The PtdIns(4,5) P_2 -induced voltage-dependent inhibition is reminiscent of the action of G proteins on N-type and P/Q-type channels^{2,16}, which has been modelled as a shift of channel gating from a 'willing' to a 'reluctant' mode on the binding of G proteins to the channel¹⁶. This model might also account for the PtdIns(4,5) P_2 inhibition. We thought that channels in the willing mode might be opened more easily by weak or moderate depolarizations than those

in the reluctant mode and that PtdIns(4,5) P_2 action might convert channels from the willing mode to the reluctant mode. According to this hypothesis, most channels would exist in the willing mode in the cell-attached configuration and immediately after patch excision (81% and 83%, respectively, $n = 37$), and a small population ($\sim 10\%$) of channels would be in the reluctant mode. Application of exogenous PtdIns(4,5) P_2 would shift most channels to the reluctant mode, reducing the fraction of willing channels from 85% to 3%, and increasing the fraction of reluctant channels from 5% to 80%.

The above results create a puzzle: if PtdIns(4,5) P_2 shifts channels to the reluctant mode, then why are most channels in intact oocytes in the willing mode? One possibility is that the endogenous PtdIns(4,5) P_2 concentration is too low to convert most channels to the reluctant mode. Another possibility is that other signalling mechanisms antagonize the inhibitory action of PtdIns(4,5) P_2 and protect the channels from PtdIns(4,5) P_2 -induced gating shift. We investigated whether PKA phosphorylation might have such a role as it regulates the interaction of PtdIns(4,5) P_2 with Kir channels¹⁷ and it produces opposite actions on Ca^{2+} channels to those induced by PtdIns(4,5) P_2 ; that is, it augments currents evoked by weak and moderate depolarizations and shifts the I - V in the negative direction^{3,18,19}.

With 2 mM Mg-ATP in the intracellular solution, PtdIns(4,5) P_2 failed to produce a substantial inhibition of $I_{\text{tail}(+10)}$ or a positive shift of the activation curve (Fig. 3a, b). This failure was observed in five of nine patches. By contrast, these PtdIns(4,5) P_2 -induced effects were observed in 25 of 25 patches without Mg-ATP. The protective action of Mg-ATP was probably due to the activation of endogenous

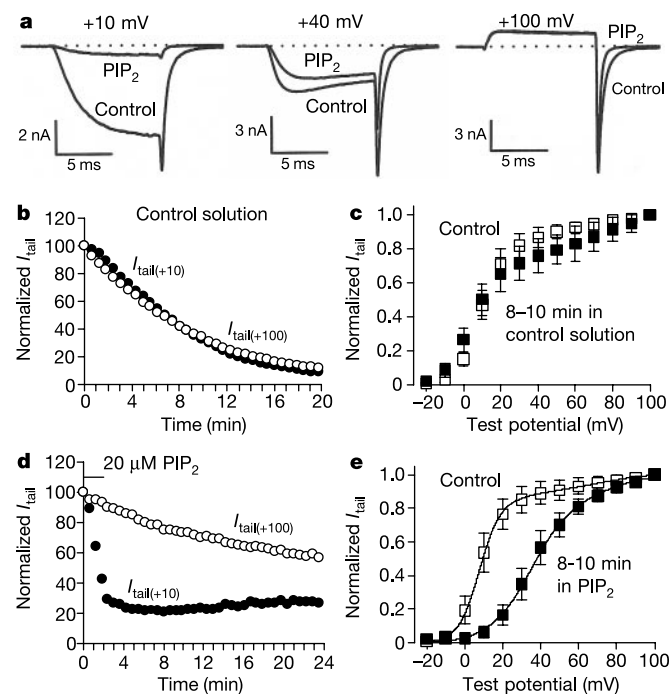


Figure 2 Voltage-dependent inhibition of P/Q-type Ca^{2+} channels by PtdIns(4,5) P_2 (PIP_2). **a**, Currents recorded immediately before and 80–100 s after application of 20 μM PtdIns(4,5) P_2 evoked by the indicated depolarization. Dotted line indicates zero current. **b**, Time course of rundown of $I_{\text{tail}(+10)}$ (filled circles) and $I_{\text{tail}(+100)}$ (open circles) in control solution. Current is normalized to that obtained immediately after patch excision. **c**, Voltage dependence of activation in control solution 1 min (open squares) and 8–10 min (filled squares) after patch excision ($n = 8$). **d**, Time course of PtdIns(4,5) P_2 action on $I_{\text{tail}(+10)}$ and $I_{\text{tail}(+100)}$. Current is normalized to that obtained before application of PtdIns(4,5) P_2 . **e**, Voltage dependence of activation before (open squares) and 8–10 min during (filled squares) application of 20 μM PtdIns(4,5) P_2 ($n = 7$). Solid lines are fitting curves (see Methods).

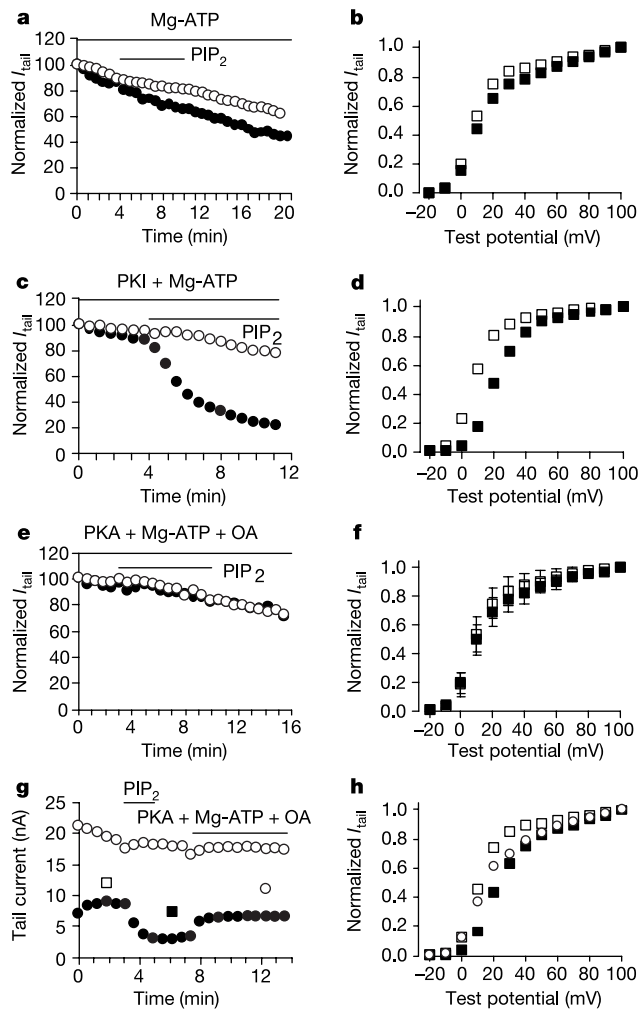


Figure 3 Crosstalk between phosphorylation by PKA and PtdIns(4,5)P₂ (PIP₂). **a, c, e**, Time course of $I_{tail(+10)}$ (filled circles) and $I_{tail(+100)}$ (open circles) in the presence of 2 mM Mg-ATP (**a**), 2 mM Mg-ATP plus 5 μ M PKI (**c**), or 2 mM Mg-ATP plus 8.8 units ml⁻¹ PKA catalytic subunit plus 0.3 μ M okadaic acid (**e**), with 20 μ M PtdIns(4,5)P₂ added during rundown. Current is normalized to that obtained immediately after patch excision. **b, d, f**, Voltage dependence of activation before (open squares) and 8–10 min after (filled squares) application of 20 μ M PtdIns(4,5)P₂. **b** and **d** are from the same patch shown in **a** and **c**, respectively, and **f** is the average of nine patches with the same experimental conditions as **e**. **g**, Time course of $I_{tail(+10)}$ (filled circles) and $I_{tail(+100)}$ (open circles) with PtdIns(4,5)P₂ treatment, followed by 2 mM Mg-ATP plus 8.8 units ml⁻¹ PKA catalytic subunit plus 0.3 μ M okadaic acid. **h**, Voltage dependence of activation obtained at the times indicated by the symbols described in **g**.

PKA in the patch because the PtdIns(4,5)P₂ effects were restored when PKI 5-24 (1–5 μ M) or H-89 (2 μ M), both of which are potent inhibitors of PKA, was added with Mg-ATP (6/7 patches for PKI 5-24 and 5/7 patches for H-89; Fig. 3c, d). When phosphorylation by PKA was strongly promoted by adding the catalytic subunit of PKA and okadaic acid together with Mg-ATP, the variability seen with Mg-ATP alone was eliminated and, in nine of nine patches, PtdIns(4,5)P₂ no longer inhibited $I_{tail(+10)}$ (Fig. 3e) or shifted the activation curve (Fig. 3f). In addition, after $I_{tail(+10)}$ was reduced by a 2-min pulse of PtdIns(4,5)P₂, subsequent application of a mixture that promoted PKA phosphorylation could quickly (albeit partially) recover the current (Fig. 3g) and shift the activation curve back towards the control curve (Fig. 3h), suggesting that phosphorylation by PKA can revert channels from the reluctant to the willing mode. By comparison, no such fast recovery was observed after washing out PtdIns(4,5)P₂ with the control solution (Fig. 2d).

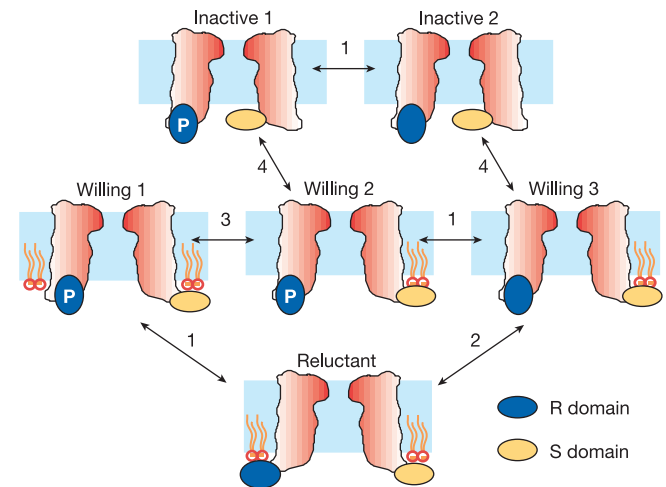


Figure 4 Model of modulation of P/Q-type Ca²⁺ channels by PtdIns(4,5)P₂ and PKA. The channel contains two distinct PtdIns(4,5)P₂-binding domains: the high-affinity S (for stabilization) domain and the low-affinity R (for reluctant and regulatory) domain. The latter can be phosphorylated by PKA, which dislodges PtdIns(4,5)P₂ from this domain. Channels without PtdIns(4,5)P₂ bound to the S domain are inactive, whereas those with and without PtdIns(4,5)P₂ on the R domain are in the reluctant and willing mode, respectively. Transitions between the various modulated states represent phosphorylation or dephosphorylation of the R domain (transition 1); binding or unbinding of PtdIns(4,5)P₂ to or from the R domain (transition 2); removal or addition of PtdIns(4,5)P₂ from or to the membrane (transition 3); and binding and unbinding of PtdIns(4,5)P₂ to or from the S domain (transition 4).

These results indicate that phosphorylation by PKA not only prevents but also reverses PtdIns(4,5)P₂-induced inhibition. We also attempted to increase endogenous phosphorylation by PKA by treating intact oocytes with forskolin or 8-bromo cyclic AMP; however, these treatments failed to produce any clear and consistent effect on P/Q-type channel currents.

The sites of PtdIns(4,5)P₂ binding and PKA phosphorylation remain to be determined. Because the Ca_v2.1 subunit contains several consensus sites for PKA phosphorylation²⁰ and PtdIns(4,5)P₂ binds directly to Kir channels¹⁰, we thought that the P/Q-type channel itself might be the target. We therefore propose a model (Fig. 4) in which P/Q-type channels possess two distinct PtdIns(4,5)P₂ binding sites: the S domain binds PtdIns(4,5)P₂ to stabilize channel activity, and the R domain binds PtdIns(4,5)P₂ to shift channels to the reluctant mode. The S domain has a higher affinity for PtdIns(4,5)P₂ than has the R domain. Finally, phosphorylation of the R domain by PKA markedly lowers its affinity for PtdIns(4,5)P₂ through electrostatic repulsion and/or allosteric conformational changes and converts channels to the willing mode.

This model can satisfactorily explain our experimental results. In intact oocytes, the concentration of endogenous PtdIns(4,5)P₂ is sufficiently high to maintain the activity of most channels by binding to the S domain. Owing to PKA phosphorylation and/or lower affinity, most channels are devoid of PtdIns(4,5)P₂ in the R domain and are thus in one of three interconvertible willing states, which we assume are biophysically indistinguishable. A very few channels that are not phosphorylated by PKA and bound to PtdIns(4,5)P₂ in the R domain are in the reluctant mode. After patch excision, the concentration of PtdIns(4,5)P₂ gradually decreases (presumably owing to dephosphorylation by lipid phosphatases), resulting in the unbinding of PtdIns(4,5)P₂ from the S domain and causing the channels to enter an inactive state (that is, rundown). Meanwhile, most channels are rapidly dephosphorylated and become susceptible to PtdIns(4,5)P₂ binding to the R domain. Thus, the application of exogenous PtdIns(4,5)P₂ on the

one hand restores PtdIns(4,5)P₂ binding to the S domain, thereby increasing the number of functional channels and stabilizing channel activity, and on the other hand converts channels to the reluctant mode. Phosphorylation by PKA interferes with PtdIns(4,5)P₂ binding to the R domain, not only preventing this conversion but also shifting reluctant channels back to the willing mode.

We next studied the effect of receptor-mediated hydrolysis of PtdIns(4,5)P₂ on whole-oocyte P/Q-type currents. To avoid complications from G-protein-induced inhibition, we examined the nerve growth factor (NGF)-activated TrkA receptor, which stimulates PtdIns(4,5)P₂ hydrolysis by directly activating phospholipase C- γ (PLC- γ)²¹. According to our model (Fig. 4), depletion of PtdIns(4,5)P₂ in intact oocytes should have two sequential functional effects: first, it should shift the small fraction (~10%) of reluctant channels to the willing mode; and second, it should induce channel downregulation. We found that NGF produces a biphasic effect on current elicited by depolarization to +10 mV ($I_{\text{peak}(+10)}$): the current increased transiently and then decreased even after washout of the agonist (Fig. 5a), probably owing to prolonged activation of TrkA receptors and/or downstream signalling pathways²². The average increase and decrease were $9 \pm 6\%$ and $34 \pm 19\%$ ($n = 31$), respectively. The I - V curve measured during the enhancement showed a slight negative shift, consistent with conversion of reluctant channels to the willing mode, and the curve measured during the inhibition was essentially a scale down, as expected from channel rundown to the inactive state (Supplementary Information).

Activation of TrkA receptors leads to stimulation of the mitogen-activated-protein-kinase (MAPK) pathway and PLC- γ ²¹, which catalyses the hydrolysis of PtdIns(4,5)P₂ into diacylglycerol (DAG) and inositol-1,4,5-triphosphate (InsP₃), two second-messengers that activate protein kinase C (PKC) and release Ca²⁺ from intracellular stores, respectively. Much evidence indicates that the NGF-induced effects resulted directly from TrkA receptor-mediated breakdown of PtdIns(4,5)P₂ rather than downstream signalling molecules and pathways. First, activation of the Tyr499Phe mutant of TrkA, which is uncoupled from the MAPK pathway^{13,21}, produced

similar effects ($8.4 \pm 3.8\%$ increase followed by $15.4 \pm 10\%$ inhibition, $n = 18$) to those of wild-type receptors. But activation of the Tyr794Phe mutant TrkA, which is unable to stimulate PLC- γ ^{13,21}, produced neither enhancement nor inhibition (Fig. 5b), indicating that hydrolysis of PtdIns(4,5)P₂ is essential for both actions. Second, preincubation of oocytes with thapsigargin, which depletes intracellular Ca²⁺ stores, abolished NGF-induced increase in intracellular concentrations of Ca²⁺ but did not affect the NGF-induced regulation of P/Q-type channels ($9.0 \pm 6.4\%$ increase, $42 \pm 16.4\%$ inhibition, $n = 10$, Fig. 5c), indicating that InsP₃-induced release of Ca²⁺ has no role. Third, activation of PKC by a phorbol ester, phorbol-12-myristate-13-acetate (PMA) (10–500 nM), produced a similar biphasic effect on P/Q-type channels as did NGF ($8 \pm 4\%$ increase, $30 \pm 18\%$ inhibition, $n = 14$). But, whereas PMA-induced regulation was completely blocked by the PKC inhibitor bisindolylmaleimide I in 17 of 22 oocytes, the NGF-induced effects remained unchanged in 12 of 14 oocytes ($8.0 \pm 4.4\%$ increase, $31 \pm 16\%$ inhibition, $n = 12$, Fig. 5d). These results indicate although PKC activation regulates P/Q-type channels, it is not required for the NGF action. Last, application of InsP₃ (100 μ M) or DAG (20 μ M) to the intracellular side in inside-out patches had no effect on $I_{\text{tail}(+10)}$ or $I_{\text{tail}(+100)}$, indicating that these second messengers do not mediate the effects of NGF (Supplementary Information).

We further investigated whether PtdIns(4,5)P₂ regulates the activity of VGCCs in native cells. In bullfrog sympathetic neurons, the neuropeptide chicken type II luteinizing hormone-releasing hormone (LHRH) induces both voltage-dependent and voltage-independent inhibition of N-type Ca²⁺ channels²³. Because stimulation of LHRH receptors activates PLC²⁴, we tested the hypothesis that the latter effect may be the result of PtdIns(4,5)P₂ breakdown. With a physiological concentration (~213 nM) of free Ca²⁺ in the pipette solution, LHRH markedly inhibited the tail current after pulses to +60 mV ($I_{\text{tail}(+60)}$) (Fig. 5e). There was usually no or only partial recovery after washout of LHRH. This inhibition was largely voltage independent (Fig. 5g). Notably, the voltage-independent component was significantly greater than that obtained with low concentrations (~10 nM) of free Ca²⁺ (ref. 23), consistent with a

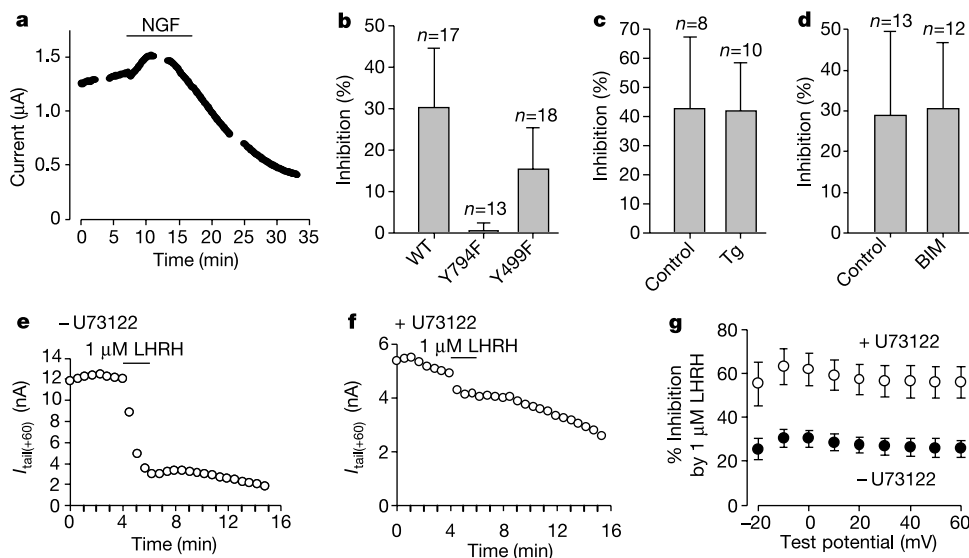


Figure 5 Regulation of P/Q- and N-type Ca²⁺ channels through receptor-mediated hydrolysis of PtdIns(4,5)P₂. **a–d**, P/Q-type channels in oocytes. **a**, Biphasic effect of NGF (100 ng ml⁻¹) on the whole-oocyte $I_{\text{peak}(+10)}$ in an oocyte expressing TrkA receptors. I - V curves were measured during the gaps. **b–d**, Comparison of NGF-induced inhibition of $I_{\text{peak}(+10)}$ in oocytes under various experimental conditions: expression of either wild-type or mutant TrkA receptors (**b**); with or without overnight thapsigargin (Tg, 100 nM)

treatment (**c**); and with or without overnight bisindolylmaleimide I (BIM, 10 μM) preincubation (**d**). **e–g**, N-type channels in bullfrog sympathetic neurons. **e**, **f**, Inhibition of $I_{\text{tail}(+60)}$ by LHRH in a neuron without (**e**) and with (**f**) U73122 preincubation (5 μM for 10 min). **g**, Comparison of LHRH-induced inhibition of the tail current after depolarization to the indicated voltages in neurons without ($n = 8$) and with ($n = 6$) U73122 pretreatment. Error bars represent s.e.m.

role for PLC, which requires Ca^{2+} for maximal activity²⁵. Most notably, the LHRH-induced inhibition was greatly attenuated in neurons pretreated with a membrane-permeable PLC inhibitor U73122 (Fig. 5f, g); however, it was not altered by pretreating the neurons with a PKC blocker staurosporin (0.5–1 μM for 30 min); LHRH reduced $I_{\text{tail} (+60)}$ by $62 \pm 11\%$ ($n = 7$) and $62 \pm 7\%$ ($n = 4$) in untreated and treated neurons, respectively. In addition, dialysing the cells with 100 μM InsP_3 neither inhibited the current nor occluded LHRH-induced inhibition ($50 \pm 7\%$, $n = 4$). These results suggest that the LHRH-induced inhibition is most probably a direct consequence of a decrease in the membrane quantity of $\text{PtdIns}(4,5)\text{P}_2$ rather than being mediated by PKC, InsP_3 or Ca^{2+} . In our model (Fig. 4), this inhibition corresponds to the conversion of active channels (predominantly in the willing mode) to inactive channels owing to loss of $\text{PtdIns}(4,5)\text{P}_2$ from the stabilization domain.

Our study shows that $\text{PtdIns}(4,5)\text{P}_2$ has two distinct and opposing actions on VGCCs: stabilization of channel activity and voltage-dependent inhibition. A further twist is that the inhibitory action can be prevented and reversed by phosphorylation by PKA. Thus, the precise modulatory action of $\text{PtdIns}(4,5)\text{P}_2$ on VGCCs depends not only on the dynamic concentration of membrane $\text{PtdIns}(4,5)\text{P}_2$, which is determined by the activities of phospholipases as well as lipid kinases and phosphatases, but also on the concentration and state of other signalling pathways. As $\text{PtdIns}(4,5)\text{P}_2$ and its regulating enzymes are present at synapses^{26–28}, and $\text{PtdIns}(4,5)\text{P}_2$ is thought to be important in exocytosis and endocytosis^{27,28}, regulation of VGCCs by $\text{PtdIns}(4,5)\text{P}_2$ could have profound consequences on synaptic transmission and plasticity.

Note added in proof: While this paper was being revised, a paper on modulation of M-current channels by $\text{PtdIns}(4,5)\text{P}_2$ appeared: Suh, B.-C. & Hille, B. Recovery from muscarinic modulation of M current channels requires phosphatidylinositol 4,5-bisphosphate synthesis. *Neuron* **35**, 507–520 (2002). □

Methods

Oocyte expression and neuron isolation

Rabbit brain $\text{Ca}_v2.1$, rat brain β_4 , rabbit skeletal muscle $\alpha_2\delta$, rat TrkA and its mutants, and rat p75 were subcloned into variants of pGEMHE or pUCHE. We synthesized cRNAs *in vitro* and injected them in various combinations into *Xenopus* oocytes, which were obtained and maintained as described²⁹.

Sympathetic ganglia were isolated from mature bullfrogs (*Rana catesbeiana*) and single neurons were isolated as described³⁰. The dissociated cells were kept in 73% L15 medium (4°C) and used within 48 h.

Electrophysiology

For two-electrode voltage-clamp studies, the electrodes were filled with 3 M KCl and had resistances of ~0.5–1 M Ω . The bath solution contained (in mM) 10 BaCl₂, 5 KCl, 60 tetraethylammonium ion (TEA)-OH, 20 NaOH and 5 HEPES (pH 7.4 with HCl). Currents were evoked every 6 s by 40-ms depolarizations to various levels from a holding potential of –80 mV. For macropatch recordings from oocytes, the pipettes had a diameter of 15–30 μm and were filled with a solution containing (in mM) 45 BaCl₂, 80 KCl and 10 HEPES (pH 7.3 with KOH). The control bath solution contained (in mM) 125 KCl, 4 NaCl, 10 HEPES and 10 EGTA (pH 7.3 with KOH). Macroscopic currents were evoked from a holding potential of –80 mV every 3 s by 10-ms depolarizations ranging from –20 mV to +100 mV in 10-mV increments, followed by a 15-ms repolarization to –20 mV. For whole-cell recordings from sympathetic neurons, the pipettes were fire-polished and had resistances of ~2–3 M Ω when filled with a solution containing (in mM) 120 CsCl, 10 HEPES, 8 EGTA, 2 MgCl₂, 5 CaCl₂, 2 Na₂-ATP and 0.5 Na-GTP (pH 7.3 with CsOH). The series resistance was compensated by 70–90%. The control bath solution contained (in mM) 2 BaCl₂, 145 TEA-Cl, 10 glucose and 10 HEPES (pH 7.3 with TEA-OH), with 2.5 μM nimodipine to block L-type channels. Currents were evoked from a holding potential of –80 mV every 3 s by 15-ms depolarizations from –50 mV to +60 mV in 10-mV increments, followed by a 15-ms repolarization to –50 mV. Currents were filtered at 2 kHz, digitized at 10–50 kHz and analysed with pClamp8 (Axon Instruments). We carried out experiments at 20–22°C.

We used the following reagents: purified brain $\text{PtdIns}(4,5)\text{P}_2$ (Avanti Polar Lipids), which was sonicated for about 15 min before application; the catalytic subunit of PKA, okadaic acid, PKI 5-24, PMA, diacylglycerol, U73122, bisindolylmaleimide I, staurosporin and H-89 (Calbiochem); antibodies against $\text{PtdIns}(4,5)\text{P}_2$ and PtdInsP (Assay Designs); Mg-ATP, thapsigargin, InsP_3 , nimodipine and trypsin (Sigma); NGF and dispace II (Boehringer Mannheim); collagenase type I (Worthington Biochemical); and chicken type II LHRH (Peninsula Laboratories). Stock solutions of water-insoluble reagents were made

in dimethyl sulphoxide, stored at –20 °C and diluted to appropriate final concentrations before use.

Data analysis

The voltage-dependence of channel activation was fitted with the sum of three Boltzmann functions of the form $f_1/1 + \exp[-(V - V_1)/k_1] + f_2/1 + \exp[-(V - V_2)/k_2] + f_3/1 + \exp[-(V - V_3)/k_3]$, where f_x , V_x and k_x are, respectively, the fraction, midpoint of activation and slope factor of the three components, with the constraint of $f_1 + f_2 + f_3 = 1$. The first two functions describe the willing and reluctant channels, respectively, with parameters $V_1 = 7.7$ mV, $k_1 = 5.5$ mV, $V_2 = 35$ mV and $k_2 = 10$ mV, which were chosen by fitting the activation curves, obtained either in cell-attached recordings or 8–10 min after application of 20 μM $\text{PtdIns}(4,5)\text{P}_2$, with the sum of two Boltzmann components. The third function has no theoretical meaning but is required to correct for voltage-dependent inactivation, which is most pronounced for inward currents evoked by depolarizations between +30 and +70 mV. V_3 and k_3 are 73 mV and 13.7 mV, respectively, on the basis of fitting the activation curve from patches that showed relatively strong inactivation. Thus, in the final fitting process, the only variables were the fractions of the three components.

T_{50} was defined as the time for current to decay to 50% of its starting value. In cases in which the recording was lost before T_{50} was reached, the decay phase of the current was fitted with either a single or double exponential function and T_{50} was extrapolated on the basis of the fitting. These patches were held for more than 15 min after patch excision and thus T_{50} was more than 15 min. About 50–70% of the patches under each condition (except the control) in Fig. 1c were analysed by this method. To verify the validity of such analysis, we intentionally truncated the data to 15 min from two 30-min recordings and found that the extrapolated T_{50} was nearly identical to the measured value. In experiments to study the effect of NGF on P/Q-type channels, NGF (100 ng ml^{–1}) was applied for 9 min. The effect was quantified according to the equation $I_x/I_{\text{Con}} = 100 \cdot [(I_{\text{Con}} - I_x)/I_{\text{Con}}]$, where I_{Con} is the current immediately before NGF application and I_x is the maximum current during NGF application (enhancement) or the current obtained 9 min after washout of NGF (inhibition).

Unless indicated otherwise, data are presented as the mean $\pm$ s.d., with the number of observations in parentheses.

Received 8 July; accepted 20 August 2002; doi:10.1038/nature01118.

- Hille, B. *Ion Channels of Excitable Membranes* (Sinauer, Sunderland, 2001).
- Catterall, W. A. Structure and regulation of voltage-gated Ca^{2+} channels. *Annu. Rev. Cell Dev. Biol.* **16**, 521–555 (2000).
- McDonald, T. F., Pelzer, S., Trautwein, W. & Pelzer, D. J. Regulation and modulation of calcium channels in cardiac, skeletal, and smooth muscle cells. *Physiol. Rev.* **74**, 365–507 (1994).
- Chad, J. E. & Eckert, R. An enzymatic mechanism for calcium current inactivation in dialysed *Helix* neurons. *J. Physiol. (Lond.)* **378**, 31–51 (1986).
- Armstrong, D. & Eckert, R. Voltage-activated calcium channels that must be phosphorylated to respond to membrane depolarization. *Proc. Natl Acad. Sci. USA* **84**, 2518–2522 (1987).
- Hao, L. Y., Kameyama, A. & Kameyama, M. A cytoplasmic factor, calpastatin and ATP together reverse run-down of Ca^{2+} channel activity in guinea-pig heart. *J. Physiol. (Lond.)* **514**, 687–699 (1999).
- Hilgemann, D. W. & Ball, R. Regulation of cardiac Na^+ , Ca^{2+} exchange and K_{ATP} potassium channels by PIP_2 . *Science* **273**, 956–959 (1996).
- Hilgemann, D. W. Cytoplasmic ATP-dependent regulation of ion transporters and channels: mechanisms and messengers. *Annu. Rev. Physiol.* **59**, 193–220 (1997).
- Hilgemann, D. W., Feng, S. & Nasuhoglu, C. The complex and intriguing lives of PIP_2 with ion channels and transporters. *Science STKE* [online] (http://stke.sciencemag.org/cgi/content/full/OC_sigtrans2001/111/re19) (2001).
- Huang, C.-L., Feng, S. & Hilgemann, D. W. Direct activation of inward rectifier potassium channels by PIP_2 and its stabilization by $\text{G}_{\beta\gamma}$. *Nature* **391**, 803–806 (1997).
- Womack, K. B. *et al.* Do phosphatidylinositides modulate vertebrate phototransduction? *J. Neurosci.* **20**, 2792–2799 (2000).
- Kobrinshi, E., Mirshahi, T., Zhang, H., Jin, T. & Logothetis, D. E. Receptor-mediated hydrolysis of plasma membrane messenger PIP_2 leads to K^+ -current desensitization. *Nature Cell Biol.* **2**, 507–514 (2000).
- Chuang, H.-h. *et al.* Bradykinin and nerve growth factor release the capsaicin receptor from $\text{PtdIns}(4,5)\text{P}_2$ -mediated inhibition. *Nature* **411**, 957–962 (2001).
- Runnels, L. W., Yue, L. & Clapham, D. E. The TRPM7 channel is inactivated by PIP_2 hydrolysis. *Nature Cell Biol.* **4**, 329–336 (2002).
- Fukami, K. *et al.* Antibody to phosphatidylinositol 4,5-bisphosphate inhibits oncogene-induced mitogenesis. *Proc. Natl Acad. Sci. USA* **85**, 9057–9061 (1988).
- Bean, B. P. Neurotransmitter inhibition of neuronal calcium currents by changes in channel voltage dependence. *Nature* **340**, 153–156 (1989).
- Liou, H.-H., Zhou, S.-S. & Huang, C.-L. Regulation of ROMK1 channel by protein kinase A via a phosphatidylinositol 4,5-bisphosphate-dependent mechanism. *Proc. Natl Acad. Sci. USA* **96**, 5820–5825 (1999).
- Bean, B. P. *Ion Channels in the Cardiovascular System; Function and Dysfunction* (eds Spooner, P. M., Brown, A. M., Catterall, W. A., Kaczorowski, G. J. & Strauss, H. C.) 237–252 (Futura, New York, 1994).
- Bean, B. P., Nowicky, M. C. & Tsien, R. W. β -Adrenergic modulation of calcium channels in frog ventricular heart cells. *Nature* **307**, 371–375 (1984).
- Mori, Y. *et al.* Primary structure and functional expression from complementary DNA of a brain calcium channel. *Nature* **350**, 398–402 (1991).
- Stephens, R. *et al.* Trk receptors use redundant signal transduction pathways involving SHC and PLC- γ 1 to mediate NGF responses. *Neuron* **12**, 691–705 (1994).
- Choi, D.-Y., Toledo-Aral, J. J., Segal, R. & Haleboua, S. Sustained signaling by phospholipase C- γ mediates nerve growth factor-triggered gene expression. *Mol. Cell Biol.* **21**, 2695–2705 (2001).
- Boland, L. M. & Bean, B. P. Modulation of N-type calcium channels in bullfrog sympathetic neurons by luteinizing hormone-releasing hormone: kinetics and voltage dependence. *J. Neurosci.* **12**, 516–533 (1993).

24. McArde, C. A., Franklin, J., Green, L. & Hislop, J. N. Signaling, cycling and desensitization of gonadotrophin-releasing hormone receptors. *J. Endocrinol.* **173**, 1–11 (2002).
25. Rhee, S. G. Regulation of phosphoinositides-specific phospholipase C. *Annu. Rev. Biochem.* **70**, 281–312 (2001).
26. McPherson, P. S. *et al.* A presynaptic inositol-5-phosphatase. *Nature* **379**, 353–357 (1996).
27. Wenk, M. R. *et al.* PIP kinase $\text{I}\gamma$ is the major PI(4,5) P_2 synthesizing enzyme at the synapse. *Neuron* **32**, 79–88 (2001).
28. Cremona, O. & De Camilli, P. Phosphoinositides in membrane traffic at the synapse. *J. Cell Sci.* **114**, 1041–1052 (2001).
29. Lu, T., Nguyen, B., Zhang, X.-M. & Yang, J. Architecture of a K^+ channel inner pore revealed by stoichiometric covalent modification. *Neuron* **22**, 571–580 (1999).
30. Yang, J. & Tsien, R. W. Enhancement of N- and L-type calcium channel currents by protein kinase C in frog sympathetic neurons. *Neuron* **10**, 127–136 (1993).

Supplementary Information accompanies the paper on Nature's website
(<http://www.nature.com/nature>).

Acknowledgements We thank S. Siegelbaum for comments on the manuscript; Y. Mori for $\text{Ca}_v2.1$ cDNA; E. Perez-Reyes for $\beta 4$ cDNA; T. Tanabe for $\alpha 2\delta$ cDNA; D. J. Julius for p75 and TrkA (wild-type and mutant) cDNAs. This work was supported by a grant to J.Y. from the National Heart, Lung, and Blood Institute. J.Y. is a recipient of the McKnight Scholar Award and the Scholar Research Programme of the EJLB Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.Y.
(e-mail: jy160@columbia.edu).

Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression

Wade Winkler*, Ali Nahvi† & Ronald R. Breaker*

* Department of Molecular, Cellular and Developmental Biology, and

† Department of Molecular Biophysics and Biochemistry, Yale University, PO Box 208103, New Haven, Connecticut 06520-8103, USA

Although proteins fulfil most of the requirements that biology has for structural and functional components such as enzymes and receptors, RNA can also serve in these capacities. For example, RNA has sufficient structural plasticity to form ribozyme^{1,2} and receptor^{3,4} elements that exhibit considerable enzymatic power and binding specificity. Moreover, these activities can be combined to create allosteric ribozymes^{5,6} that are modulated by effector molecules. It has also been proposed^{7–12} that certain messenger RNAs might use allosteric mechanisms to mediate regulatory responses depending on specific metabolites. We report here that mRNAs encoding enzymes involved in thiamine (vitamin B₁) biosynthesis in *Escherichia coli* can bind thiamine or its pyrophosphate derivative without the need for protein cofactors. The mRNA–effector complex adopts a distinct structure that sequesters the ribosome-binding site and leads to a reduction in gene expression. This metabolite-sensing regulatory system provides an example of a ‘riboswitch’ whose evolutionary origin might pre-date the emergence of proteins.

A thiamine pyrophosphate (TPP)-dependent sensor/regulatory protein has been proposed to participate in the control of thiamine biosynthetic genes¹³, but thus far no such protein factor has been shown to exist. Recently, we established that the mRNA leader sequence of the *btuB* gene of *E. coli* can bind coenzyme B₁₂ selectively, and that this binding event brings about a structural change in the RNA that is important for genetic control¹⁴. In the current study, we set out to determine whether mRNAs that encode thiamine biosynthetic proteins might also use a riboswitch that is

located in their untranslated leader sequence. We prepared β -galactosidase translational fusion constructs that encompass the 5'-untranslated regions of *thiM* and *thiC* mRNAs of *E. coli*, each of which includes a previously identified ‘*thi* box’ domain whose sequence and potential secondary structure are conserved in several species of bacteria and archaea¹¹. The *thiM* and *thiC* translational fusion constructs exhibit thiamine-dependent suppression of β -galactosidase activity of 18- and 110-fold, respectively, when host cells are grown in a minimal medium. A transcriptional fusion containing the *thiM* leader is not subject to suppression by thiamine, but a similar fusion with *thiC* leader yields a 16-fold modulation with thiamine, suggesting that a significant portion of genetic control observed with *thiC* occurs at the level of transcription.

These constructs were subsequently used to prepare DNA templates by polymerase chain reaction (PCR) for *in vitro* transcription of RNA fragments. The resulting RNAs were subjected to a structure-probing process^{14–16} to reveal whether the RNAs undergo structure modulation upon binding of ligands. Internucleotide linkages in unstructured regions are more likely to undergo spontaneous cleavage compared to linkages that reside in highly structured regions of an RNA¹⁵. The 165-nucleotide *thiM* RNA fragment (165 *thiM*) has a distinct pattern of cleavage products that is generated when the RNA is incubated for an extended period in the absence of TPP (Fig. 1a). Upon addition of 100 μM TPP, 165 *thiM* undergoes substantial structural alteration, with many internucleotide linkages in the region spanning positions 39–80 exhibiting a reduction in spontaneous cleavage. This indicates that TPP binds to the RNA and stabilizes a defined structure within this region, resulting in a lower rate of fragmentation.

The fragmentation patterns are largely congruent with potential stem and bulge structures that are identified by a secondary-structure prediction algorithm^{17,18}. Most linkages that experience a ligand-induced reduction of cleavage are encompassed by the *thi* box and nucleotides that reside immediately 5' relative to this domain (Fig. 1b). Other linkages that undergo cleavage, but that are not modulated by TPP, are predicted to reside in bulges or in the loops of hairpins. Predicted base-paired structures labelled P2–P7 include linkages that exhibit the lowest levels of spontaneous cleavage, implying that they remain structured in both the presence and absence of TPP. Nucleotides 126–130 encompass the only region apart from those described above that becomes more structured upon TPP addition. These nucleotides correspond to the Shine–Dalgarno (SD) sequence, which is required to be unpaired for efficient translation of mRNAs in prokaryotes. These findings are consistent with a genetic control mechanism in which the *thiM* RNA binds to TPP and forms a complex such that the ribosome cannot gain access to the SD sequence.

Similarly, structure probing was used to examine the mRNA leader for *thiC*. The 240 *thiC* RNA also exhibits extensive modulation of its pattern of spontaneous cleavage, and again the majority of the changing pattern is located in the *thi* box and in the region located immediately upstream of this domain (Fig. 1c). These regions of highest structure modulation in *thiM* and *thiC* can be folded into similar secondary structures (Fig. 1d), and carry several common sequence elements within and adjacent to the *thi* box domain. Therefore, we propose that the structures of *thiM* and *thiC* spanning stems P1–P5 comprise TPP-binding motifs that are analogous to aptamers, which are engineered ligand-binding RNAs^{3,4,19}. Nucleotides residing 3' relative to this natural TPP aptamer are probably involved in converting the metabolite binding event into a genetic response (see below).

The sensitivity of metabolite detection by these mRNAs was assessed by establishing apparent dissociation constant (apparent K_D) values for TPP, thiamine, and thiamine monophosphate (TP). Values were generated by monitoring the extent of spontaneous cleavage at several ligand-sensitive sites within the RNA over a range

of ligand concentrations. For example, probing of a trace amount of 165 *thiM* RNA with TPP concentrations ranging from zero to 100 μ M (or up to 10 mM with certain analogues) reveals that half-maximal modulation of RNA structure occurs when approximately 600 nM TPP is present (Fig. 2a), which reflects an apparent K_D of 600 nM. Likewise, probing of 240 *thiC* reveals an apparent K_D of 100 nM. Both 165 *thiM* and 240 *thiC* RNAs appear to bind TPP more readily than TP or thiamine, with *thiC* exhibiting more than 1,000-fold discrimination against TP and thiamine (Fig. 2b). The fact that TPP is the strongest modulator of RNA structure is consistent with genetic observations in *Salmonella typhimurium* that TPP synthesis is required for regulation of expression of thiamine biosynthesis genes²⁰. The differential specificity achieved by the RNAs, which is a phenomenon that is commonly observed for receptor–ligand systems made of protein, indicates that these ligand-binding RNAs might be amenable to specificity changes through evolutionary forces.

It is important to note that the actual K_D values for RNA–ligand interactions might be different inside cells where physiological conditions of Mg^{2+} and other agents that can influence RNA structure will not match those of our assays. Also, the nature of the RNA construct can be a source of an altered K_D . For example, we found that the minimized 91 *thiM* construct (Fig. 1a), which largely encompasses only the putative natural aptamer, retains the ability to bind TPP and exhibits an apparent K_D that is improved by ~ 20 fold compared to the longer construct (Fig. 2b). Thus, the affinity for TPP might vary as the nascent RNA transcript emerges from the active site of RNA polymerase or the ribosome. Furthermore, this

result demonstrates that the 91 *thiM* aptamer can be separated from RNA components (collectively termed the ‘expression platform’) that are directly controlling gene expression. This modular construction, involving the physical and functional separation of aptamer and expression platform domains, would be expected to facilitate the generation of TPP-controlled RNAs through evolutionary processes or by rational RNA engineering strategies.

We noticed that spontaneous cleavage at several linkages within the *thi* box domain of 165 *thiM* specifically correlate with the type of ligand used. Although TPP reduces spontaneous cleavage of 165 *thiM* at nucleotides A61, U62 and, to a smaller extent, at U79, these same sites retain an elevated level of cleavage when thiamine is present near its saturating concentration (Fig. 2c). These nucleotides cluster at an internal bulge within the *thi* box domain, and could be contributing to the binding site for the phosphate groups of TPP.

To further explore the molecular recognition characteristics of the TPP riboswitch, we examined the structural modulation of 165 *thiM* in the presence of several analogues that carry certain structural features of thiamine (Fig. 3a). Thiamine, TP and TPP induce modulation as expected (Fig. 3b). However, oxythiamine and other thiamine analogues with less similarity to TPP fail to induce structure modulation. The performance of this sampling of analogues indicates that the RNA makes specific contacts to distal parts of its ligand, and that both the purine and phosphate groups carry important elements for molecular recognition (Fig. 3c). Similar results are obtained by using equilibrium dialysis assays (Fig. 3d). For example, the addition of 91 *thiM* RNA to chamber *b* of an

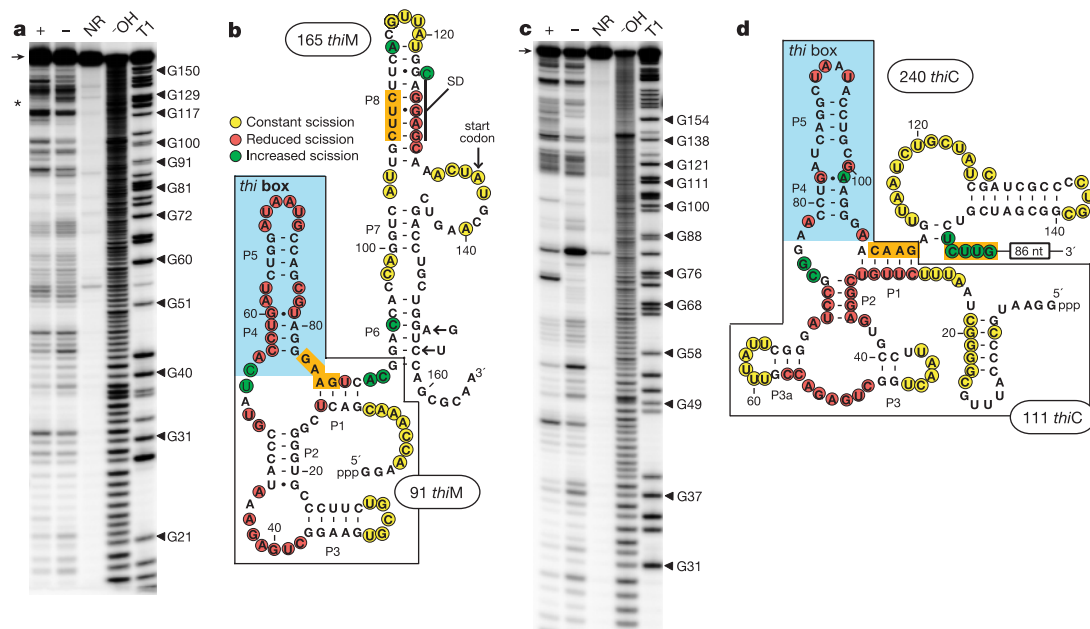


Figure 1 Metabolite binding by mRNAs. **a**, Thiamine pyrophosphate (TPP)-dependent modulation of the spontaneous cleavage of 165 *thiM* RNA visualized by polyacrylamide gel electrophoresis (PAGE). 5' ³²P-labelled RNAs (arrow, 20 nM) were incubated for ~ 40 h at 25 °C in 20 mM $MgCl_2$, 100 mM KCl, and 50 mM Tris–HCl (pH 8.3 at 25 °C) in the presence (+) or absence (–) of 100 μ M TPP. NR, OH and T1 represent RNAs subjected to no reaction, partial digestion with alkali, or partial digestion with RNase T1 (G-specific cleavage), respectively. Product bands corresponding to cleavage after selected G residues are numbered and identified by filled arrowheads. The asterisk identifies modulation of RNA structure involving the Shine–Dalgarno (SD) sequence. Gel separations were analysed using a phosphorimager (Molecular Dynamics) and quantified using ImageQuant software. **b**, Secondary-structure model of 165 *thiM* as predicted by computer modelling^{17,18} and by the structure-probing data depicted in **a**. Unmarked

nucleotides exhibit a constant but low level of degradation. The truncated 91 *thiM* RNA is boxed, and the *thi* box element¹¹ is shaded light blue. Nucleotides highlighted in orange identify a possible alternative pairing, designated P8*. The RNA carries two mutations (G155A and U157C) relative to wild type that were introduced in a non-essential portion of the construct to form a restriction site for cloning, while all RNAs carry two 5' -terminal G residues to facilitate *in vitro* transcription. **c**, TPP-dependent modulation of the spontaneous cleavage of 240 *thiC* RNA. Reactions were conducted and analysed as described in **a**. **d**, Secondary-structure model of 240 *thiC*. Potential base-paired elements that are similar to those of *thiM* are labelled P1–P5. The truncated RNA 111 *thiC* is boxed. Nucleotides highlighted in orange identify a possible alternative pairing. Additional details are as described in **b**.

equilibrium dialysis assembly (see Methods) causes a shift in the distribution of ^3H -thiamine in favour of chamber *b*, unless an excess of unlabelled TPP is also included. However, the presence of oxythiamine does not significantly restore the tritium distribution to unity, which is expected because probing data indicate that it is not able to bind the RNA. These findings imply that the aptamer domain of the TPP riboswitch is highly selective for its target ligand.

The secondary-structure model for 165 *thiM* RNA was examined in greater detail by generating and testing a series of variant constructs (Fig. 4a). For example, variant M1 carries a mutation that disrupts the predicted P3 pairing element. This mutation causes a loss of TPP binding (Fig. 4b; for example, see position C77) and a loss of genetic control of the corresponding β -galactosidase fusion construct (Fig. 4b, bottom). Restoration of base pairing in the double-mutant construct M2 restores both TPP binding and genetic control. Similarly, disruptive and restorative mutations encompassed by constructs M3–M6 are consistent with the formation of stems P5 and P8 (Fig. 4c). Upon the addition of TPP, the SD element of both the wild-type and M2 constructs

becomes sequestered in a structure that precludes a high level of spontaneous cleavage. In contrast, the M1 construct does not exhibit SD modulation (Fig. 4b, nucleotides 126–130). These results suggest that the genetic switch might be turned off by a mechanism whereby TPP binding ultimately promotes the stable formation of P8, which reduces access to the SD by the ribosome.

We observed that the predicted partner of the SD sequence in P8 (nucleotides 108–111) remains resistant to spontaneous cleavage both in the presence and absence of TPP (Fig. 1a). This suggests that, upon addition of TPP, the formation of P8 might be due to the displacement of an alternative structure that otherwise prevents this anti-SD element from forming P8. Furthermore, we noticed that nucleotides 83–86 are complementary to the anti-SD element, and that this region also resists spontaneous cleavage in the presence and absence of TPP. One mechanism by which genetic control could result is by way of the mutually exclusive formation of P8* (Fig. 1b) in the 'on' state versus the simultaneous formation of P1 and P8 in

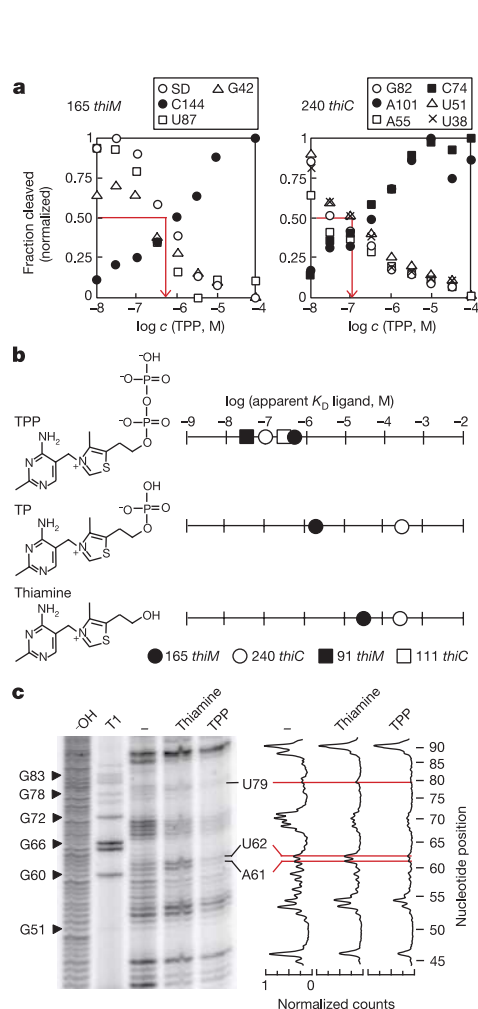


Figure 2 The *thiM* and *thiC* mRNA leaders serve as high-affinity metabolite receptors.

a, The extent of spontaneous RNA cleavage at several sites within 165 *thiM* (left) and 240 *thiC* (right) are plotted for different concentrations (*c*) of TPP. Red arrows reflect the estimated concentration of TPP needed to attain half maximal modulation of RNA (apparent K_D). **b**, The logarithm of the apparent K_D values are plotted for both RNAs with TPP, TP (thiamine monophosphate) and thiamine as indicated. **c**, Patterns of spontaneous cleavage of 165 *thiM* differ between thiamine and TPP ligands as depicted by PAGE analysis (left) and as reflected by graphs (right) representing the relative phosphorimager counts for the three lanes as indicated. Peaks corresponding to the identified bands are highlighted by coloured lines. Details for the RNA probing analysis are similar to those described in Fig. 1a. The graphs were generated by ImageQuant software.

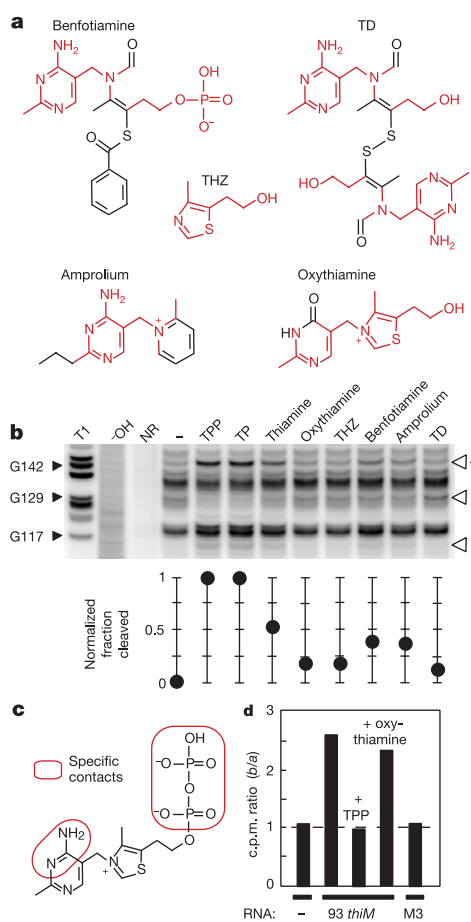


Figure 3 High sensitivity and selectivity of mRNA leaders for metabolite binding.

a, Chemical structures of several analogues of thiamine. (TD, thiamine disulphide; THZ, 4-methyl-5- β -hydroxyethylthiazole.) Structures in red are similar to TPP. **b**, PAGE analysis of 165 *thiM* RNA structure probing using TPP and various chemical analogues (40 μM each) as indicated. Locations of significant structural modulation within the RNA spanning nucleotides ~113 to ~150 are indicated by open arrowheads. The asterisk identifies the site (C144) used to compare the normalized fraction of RNA that is cleaved (bottom) in the presence of specific compounds. Details for the RNA probing analysis are similar to those described in Fig. 1a. **c**, Summary of the features of TPP that are critical for molecular recognition. **d**, Equilibrium dialysis using ^3H -thiamine as a tracer. Plotted are the ratios for tritium distribution in a two-chamber system (*a* and *b*) that were established upon equilibration in the presence of the RNA constructs in chamber *b* as indicated (see Fig. 4 for a description of the non-TPP-binding mutant M3). 100 μM TPP or oxythiamine was added to chamber *a*, as denoted, upon the start of equilibration.

the metabolite-bound 'off' state (Fig. 5). A genetic control strategy of alternative structure formation would be similar to that found in mRNAs that use terminator, antiterminator, and anti-antiterminator structural elements.

In a preliminary assessment of this mechanism, we tested constructs M7–M9. Construct M7 carries a U109C mutation in the anti-SD sequence that is expected to destabilize the P8 interaction while simultaneously destabilizing the P8* interaction. M7 retains TPP binding function and exhibits a significant level of genetic modulation (Fig. 4c), which is expected if the mutation does not disrupt the relative distribution of mRNAs between the 'on' and 'off' states. In comparison, M8 (U110C) retains TPP binding, exhibits a marked reduction in the level of reporter expression, and loses nearly all genetic modulation. In addition, M8 no longer exhibits detectable spontaneous cleavage in the SD sequence, which is consistent with the prediction that the thermodynamic balance between P8 and P8* formation should be shifted decidedly in favour of P8 in this RNA variant. Construct M9, which carries four mutations in the anti-SD element, has a significantly different pattern of spontaneous cleavage in the SD region. Not surprisingly, M9 fails to reduce gene expression upon thiamine addition to cells, despite the fact that the construct retains TPP binding activity *in vitro*. Although a more detailed analysis is needed to confirm this model, it is evident from these data that TPP binding restricts the structural freedom of the SD element in the appropriate RNA variants, and that this correlates with genetic control.

The biochemical and genetic data reported here show that certain mRNAs carry natural aptamer domains, and that binding of specific metabolites directly to these RNA domains leads to modulation of gene expression. RNA structure-probing data indicate that the TPP riboswitch operates as an allosteric sensor of its target compound—binding of TPP by the aptamer domain stabilizes a conformational state within the aptamer and within the neighbouring expression platform that precludes translation. The diversity of expression platforms seems to be large. Our findings suggest that the *thiM* RNA uses a SD-blocking mechanism to control translation. In contrast, the *thiC* RNA controls gene expression both at transcription and

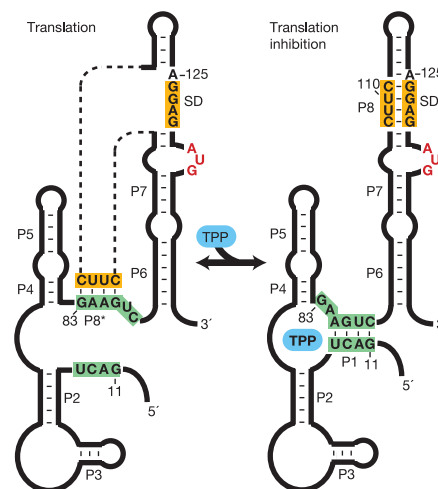


Figure 5 Schematic representation of the proposed mechanism for TPP-dependent deactivation of *thiM* translation. In the absence of TPP, the P8* pairing is formed between the anti-SD element and the anti-anti-SD element. This conformation permits the SD sequence to interact with the ribosome, and thus translation proceeds. In the presence of TPP (blue), the obligate formation of the P1 stem sequesters a portion of the anti-anti-SD element, and therefore the complete P8 stem also forms. This precludes ribosome access to the SD element, which inhibits translation. Complementary sequence elements that form P1 and P8 are depicted in green and orange, respectively.

translation, and therefore might make use of a more complex expression platform that converts TPP binding into a transcription termination event and into inhibition of translation of completed mRNAs.

In addition to the TPP-sensing system that we describe here, and a newly discovered riboswitch¹⁴ that functions with coenzyme B₁₂, we also have evidence that the *RFN* element²¹ (a conserved RNA sequence found in certain bacterial mRNAs) serves as a flavin mononucleotide-dependent riboswitch (unpublished data).

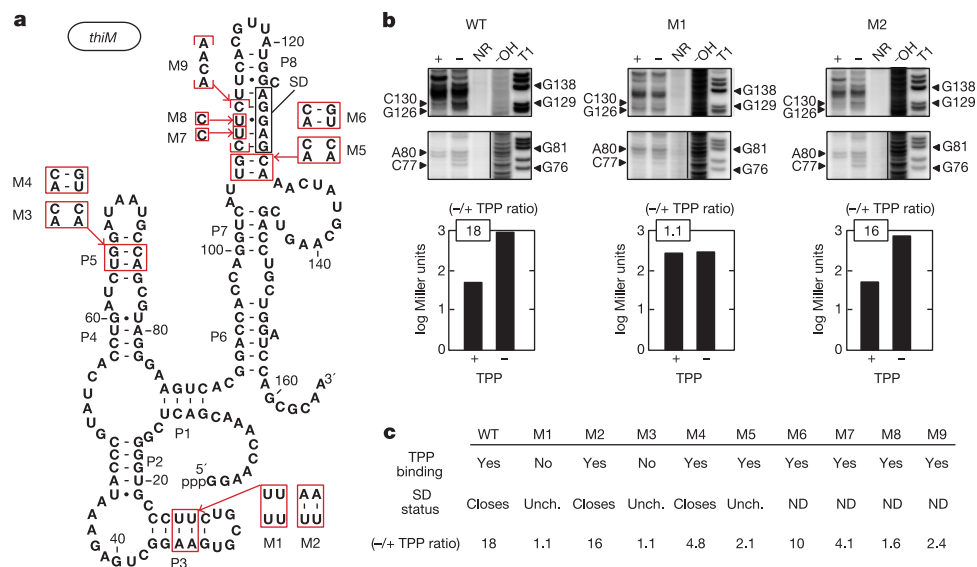


Figure 4 Mutational analysis of the structure and function of the *thiM* riboswitch. **a**, Mutations present in constructs M1–M9 relative to the 165 *thiM* RNA. Boxed nucleotides designate the locations and the identities of the mutations tested. **b**, *In vitro* ligand-binding and genetic control functions of the wild-type (WT), M1 and M2 RNAs as reflected by PAGE analysis of in-line probing experiments (10 μ M TPP) and by β -galactosidase expression assays. Labels on gels are as described in Fig. 1a. Bars represent the levels of gene expression in the presence (+) and the absence (–) of TPP in

the culture medium. **c**, A summary of analyses (as in **b**) of WT to M9 is presented in table form. The SD status 'closes' signifies that the Shine–Dalgarno sequence exhibits reduced structural flexibility (probably due to P8 formation), while the 'Unch.' designation indicates that the sequence remains structurally unchanged. The SD status 'ND' (not determined) indicates either that the level of spontaneous cleavage detected in the absence and presence of TPP is near the limit of detection (M6, M7 and M8) or that the region adopts an atypical structure (M9) compared to WT.

Although these are (to our knowledge) the first examples of mRNA elements that control genetic expression by metabolite binding, we suspect that they might be only the initial representatives of a genetic control strategy that is widespread in biology. It has been suggested^{22–24} that TPP, coenzyme B₁₂ and FMN emerged as biological cofactors during the ‘RNA world’²⁵. If these metabolites were being biosynthesized and used before the advent of proteins, then certain riboswitches might be modern examples of the most ancient form of genetic control. A preliminary search of genomic sequence databases has revealed that sequences corresponding to the TPP aptamer exist in organisms from bacteria, archaea and eukarya—largely without major alteration. Although new metabolite-binding mRNAs are likely to emerge as evolution progresses, it is possible that the known riboswitches are ‘molecular fossils’ from the RNA world. □

Methods

Chemicals and oligonucleotides

TPP, TP, thiamine, oxythiamine, amprolium and benfotiamine were purchased from Sigma. Thiamine disulphide and THZ were purchased from TCI America. ³H-labelled thiamine was purchased from American Radiolabeled Chemicals, Inc. (10 Ci mmol^{−1}). Synthetic DNAs were synthesized by the Keck Foundation Biotechnology Resource Center at Yale University. DNAs were purified by denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE), and isolated from the gel by crush-soaking in 10 mM Tris–HCl (pH 7.5 at 23 °C), 200 mM NaCl and 1 mM EDTA. The DNA was recovered by precipitation with ethanol.

Construction of *E. coli* *thiM*–*lacZ* and *thiC*–*lacZ* fusions

The region spanning nucleotides −83 to 238 of the *E. coli* *thiCEFGH* operon²⁶ was amplified by PCR from *E. coli* strain MC4100 (a gift from S. Gottesman) as an *Eco*R1–*Bgl*II fragment. The DNA was ligated into *Eco*R1- and *Bam*H1-digested pRS414 plasmid DNA, which contains a promoterless copy of *lacZ* (a gift from R. Simons; ref. 27), resulting in the in-frame fusion of the ninth codon of *lacZ* to the ninth codon of *thiC*. Similarly, the regulatory region of *thiM* (nucleotides −67 to 163) was amplified by PCR as an *Eco*R1–*Bam*H1 fragment and inserted into plasmid pRS414, wherein the sixth codon of *thiM* resides in-frame with the ninth codon of *lacZ*. The plasmids were transformed into Top10 cells (Invitrogen) for all subsequent manipulations. All site-directed mutations were introduced into the *thiC* and *thiM* regulatory regions using the QuikChange site-directed mutagenesis kit (Stratagene) and the appropriate mutagenic DNA primers. All mutations were confirmed by DNA sequencing (USB Thermosequenase).

Thiamine-repression β-galactosidase assays

E. coli cells (Top10, Invitrogen) that contained in-frame *lacZ* fusions to *thiC* or *thiM* mRNA leader sequences, were grown in M9 glucose minimal medium (plus 50 μg ml^{−1} vitamin assay Casamino acids; Difco) to mid-exponential phase. The cultures were grown either with or without added thiamine (100 μM). Aliquots (1 ml) were removed for β-galactosidase enzyme assays, which were conducted in a manner similar to that described by Miller²⁸. All experiments were repeated twice and in duplicate, with Miller unit values reflecting the average of these analyses.

In vitro transcription

Templates for *in vitro* transcription of the fragments of *thiC* and *thiM* mRNA leaders were generated by PCR using the appropriate DNA primers and plasmids pRS414*thiC* or pRS414*thiM*, respectively. The dinucleotide sequence GG was introduced into the DNA constructs (corresponding to the 5′ terminus of each RNA construct) to facilitate transcription by T7 RNA polymerase. RNAs were prepared by *in vitro* transcription and were 5′ ³²P-labelled as described previously⁶.

In-line probing of RNA

Determination of apparent *K*_D values for each construct was achieved by conducting in-line probing of RNA wherein the concentration of the ligand was varied between 10 nM and 100 μM, or up to 10 mM for weakly binding ligands (see Fig. 1a legend for details). Composite plots of the fraction of RNA cleaved at specific sites versus the logarithm of the concentration of ligand (for example, Fig. 2a) were generated to provide an estimate of the apparent *K*_D. Fraction cleaved values were normalized relative to the highest and lowest cleavage values measured for each site.

Equilibrium dialysis

Equilibrium dialysis experiments were conducted using a DispoEquilibrium Dialyzer

(ED-1, Harvard Bioscience), wherein chambers *a* and *b* were separated by a 5,000 dalton molecular mass cut-off membrane. Equilibration was initiated by the addition of 25 μl of equilibration buffer (50 mM Tris–HCl (pH 8.3 at 25 °C), 20 mM MgCl₂, 100 mM KCl), containing 100 nM ³H-thiamine, and by the addition of an equal volume of equilibration buffer either without or with 20 μM RNA as indicated to chamber *b*. Equilibrations were allowed to proceed for 10 h at 23 °C, and aliquots were removed from each chamber and quantified by using a liquid scintillation counter.

Received 2 August; accepted 20 September 2002; doi:10.1038/nature01145.

Published online 16 October 2002.

- Cech, T. R. & Golden, B. L. *The RNA World* (eds Gesteland, R. F., Cech, T. R. & Atkins, J. F.) 321–350 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1998).
- Breaker, R. R. *In vitro* selection of catalytic polynucleotides. *Chem. Rev.* **97**, 371–390 (1997).
- Osborne, S. E. & Ellington, A. D. Nucleic acid selection and the challenge of combinatorial chemistry. *Chem. Rev.* **97**, 349–370 (1997).
- Hermann, T. & Patel, D. J. Adaptive recognition by nucleic acid aptamers. *Science* **287**, 820–825 (2000).
- Soukup, G. A. & Breaker, R. R. Engineering precision RNA molecular switches. *Proc. Natl Acad. Sci. USA* **96**, 3584–3589 (1999).
- Seetharaman, S., Zivarts, M., Sudarsan, N. & Breaker, R. R. Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nature Biotechnol.* **19**, 336–341 (2001).
- Gold, L. *et al.* From oligonucleotide shapes to genomic SELEX: novel biological regulatory loops. *Proc. Natl Acad. Sci. USA* **94**, 59–64 (1997).
- Gold, L., Singer, B., He, Y. & Brody, E. SELEX and the evolution of genomes. *Curr. Opin. Gen. Dev.* **7**, 848–851 (1997).
- Nou, X. & Kadner, R. J. Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc. Natl Acad. Sci. USA* **97**, 7190–7195 (2000).
- Gelfand, M. S., Mironov, A. A., Jomantas, J., Kozlov, Y. I. & Perumov, D. A. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* **15**, 439–442 (1999).
- Miranda-Rios, J., Navarro, M. & Soberón, M. A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl Acad. Sci. USA* **98**, 9736–9741 (2001).
- Stormo, G. D. & Ji, Y. Do mRNAs act as direct sensors of small molecules to control their expression? *Proc. Natl Acad. Sci. USA* **98**, 9465–9467 (2001).
- Webb, E. & Downs, D. Characterization of *thiL*, encoding thiamin-monophosphate kinase, in *Salmonella typhimurium*. *J. Biol. Chem.* **272**, 15702–15707 (1997).
- Nahvi, A. *et al.* Genetic control by a metabolite-binding mRNA. *Chem. Biol.* **9**, 1043–1049 (2002).
- Soukup, G. A. & Breaker, R. R. Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* **5**, 1308–1325 (1999).
- Soukup, G. A., DeRose, E. E., Koizumi, M. & Breaker, R. R. Generating new ligand-binding RNAs by affinity maturation and disintegration of allosteric ribozymes. *RNA* **7**, 524–536 (2001).
- Zuker, M., Mathews, D. H. & Turner, D. H. *RNA Biochemistry and Biotechnology* (eds Barciszewski, J. & Clark, B. F. C.) 11–43 (NATO ASI Series, Kluwer Academic, Boston, 1999).
- Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940 (1999).
- Gold, L., Poliski, B., Uhlenbeck, O. & Yarus, M. Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**, 763–797 (1995).
- Webb, E., Febres, F. & Downs, D. M. Thiamine pyrophosphate (TPP) negatively regulates transcription of some *thi* genes of *Salmonella typhimurium*. *J. Bacteriol.* **178**, 2533–2538 (1996).
- Gelfand, M. S., Mironov, A. A., Jomantas, J., Kozlov, Y. I. & Perumov, D. A. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* **15**, 439–442 (1999).
- White, H. B. III Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* **7**, 101–104 (1976).
- White, H. B. III *The Pyridine Nucleotide Coenzymes* 1–17 (Academic, New York, 1982).
- Benner, S. A., Ellington, A. D. & Tauer, A. Modern metabolism as a palimpsest of the RNA world. *Proc. Natl Acad. Sci. USA* **86**, 7054–7058 (1989).
- Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
- Vander Horn, P. B., Backstrom, A. D., Stewart, V. & Begley, T. P. Structural genes for thiamine biosynthetic enzymes (*thiCEFGH*) in *Escherichia coli* K-12. *J. Bacteriol.* **175**, 982–992 (1993).
- Simons, R. W., Houman, F. & Kleckner, N. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* **53**, 85–96 (1987).
- Miller, J. H. *A Short Course in Bacterial Genetics* 72 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992).

Acknowledgements We thank members of the Breaker laboratory for comments on the manuscript, especially N. Sudarsan for discussions. This work was supported by the NIH and the NSF, and by a fellowship to R.R.B. from the David and Lucile Packard Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.R.B. (e-mail: ronald.breaker@yale.edu).

Neuroscience update

The latest products from antibodies to vision systems.

Sigma catalogue

Sigma-RBI www.sigma-aldrich.com

Hit the headlines

This 32-page brochure *Alzheimer's Disease and other Neurodegenerative Disorders* features 250 items, including 75 new products in areas such as secretase inhibitors, apolipoproteins, amyloid peptides, tau proteins and antibodies to β -amyloid, tau, α -synuclein and prion protein. The brochure also contains detailed overviews of current information on Alzheimer's, Parkinson's and other neurodegenerative conditions.

Ephrin antibodies

Zymed www.zymed.com

Ligands for ephrin receptors

Ephrin antibodies for the study of these ligands and their corresponding tyrosine kinase ephrin receptors are available from Zymed Laboratories. The company offers five rabbit polyclonal ephrin antibodies: ephrins A1, A2, A4, B1 and B2. Applications include western blotting, IHC for formalin-fixed, paraffin-embedded tissues and immunoprecipitation. A new monoclonal antibody to chicken ephrin B1 can be used for ELISAs and western blotting.

Microarrays

PerkinElmer www.perkinelmer.com/lifesciences

Maintain your profile

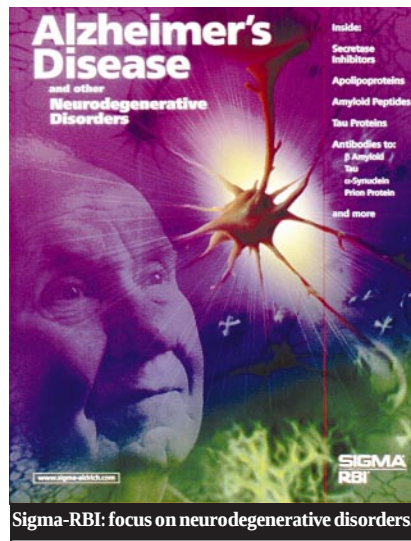
Two new glass-based microarrays have been added to the PerkinElmer range, one for apoptosis (programmed cell death) and one for neurobiology applications. The human apoptosis (MPS606) and human neurobiology (MPS607) microarrays are additions to the MICROMAX microarray product family. The apoptosis microarray contains 333 genes plus 27 control genes, and the neurobiology microarray contains 402 neurobiology genes plus 32 control genes, each represented in duplicate. Each kit contains two microarrays and a technical data sheet.

PatchMan NP 2

Eppendorf www.eppendorf.com

Fit for the purpose

The PatchMan NP2 micromanipulator from Eppendorf is designed to perform electrophysiological measurements on cells and pieces of membrane. It enables researchers to place measurement electrodes both on and in the cell. The electronic engagement mechanism on the capillary allows several hours of drift-free operation. The joystick can be



deactivated as a precaution against inadvertent movement. Other features include optional pure down-movement, rapid change of electrode and an automatic home function. The device is also suitable for use as a universal micromanipulator.

MED64

Harvard Bioscience

www.harvardapparatus.com

Wizardry

The MED64 system is an integrated, 64-electrode instrument used to perform stimulation, data acquisition and analysis in both *in vitro* and *ex vivo* samples. MED64 Performer software guides researchers through experimental protocols using wizards to streamline experiments and improve accuracy and consistency of results. Saving a completed experiment creates a file that can be distributed and analysed using the Performer Viewer, which is available free at www.med64.com. New protocols for rhythmic activity and spike sorting are

used to better understand the biological clock and network communications in the brain. Applications include drug and toxin screening, developmental electrophysiology, network physiology and patterned stimulation.

Vision system

Affinity BioReagents www.bioreagents.com

The eyes have it

Affinity BioReagents offers nine antibodies that recognize key vision system proteins: arrestin, PDE6 α , PDE6 α' , PDE6 γ , transducin rod α , transducin rod β , CaM kinase I, rhodopsin kinase 1a and rhodopsin kinase 1a/1b. These highly specific antibodies are designed to help researchers elucidate important elements of the biochemical cascade that leads to sight. They are optimized for a range of applications, including western blots, immunoprecipitations, immunohistochemistry and immunofluorescence.

CellMap

Bio-Rad www.cellscience.bio-rad.com

Be more specific

The two new confocal laser scanning microscopy systems from the Cell Science Division of Bio-Rad are designed for specific cell analyses. CellMap IC is designed for characterizing protein location in cells relative to nuclear and cellular structure, and CellMap ID is designed for determining protein location in relation to the cell cycle and for characterizing tissue morphology. Both CellMap IC and CellMap ID are to make their debuts at Neuroscience 2002 (2–7 November in Orlando), where Bio-Rad will be giving away CellMap T-shirts on the mornings of the first two days.

These notes are compiled in the Nature office from information provided by the manufacturers. Press releases to be considered for publication can be sent to newproducts@nature.com.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building

4 Crinan Street

London N1 9XW, UK

Tel +44 (0) 20 7843 4961

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Sille Opstrup (4994)

France/ Belgium:

Amelie Pequignot (4974)

Production Manager:

 Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Switzerland:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager:

 Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikatamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

e-mail: k.johnson@naturejp.com

Asia-Pacific Advertising Manager:

Kevyn Johnson

naturejobs

Job insecurity

This autumn, companies seemed to have shed jobs at the same rate as the trees lost their leaves. The cuts and consolidation don't seem to be limited to any one country or speciality — not that that is any consolation. The Icelandic genomics firm deCODE Genetics, for example, cut its worldwide workforce by 30%. Oxford GlycoSciences, the UK-based proteomics company, slashed its workforce by a fifth. In the United States, Entremed in Rockville, Maryland, has more than halved its number of staff. And one of the remaining independent bioinformatics firms of any size, Informax in Rockville, is being bought by Invitrogen — job losses are expected.

What is going on? The combination of a clamp-down on investment by venture capitalists and a declining stock market — both of which have been sustained for over a year — is draining companies who don't yet have products ready for the market, or whose business plans didn't pan out.

And more job losses are likely. Biotech analysts in the United States and Europe have long been warning that more consolidation needs to happen — there are simply too many small-to-medium-sized companies competing for the same financial resources.

And many pharmaceutical firms, concerned about their dwindling product pipelines and declining stocks, are eyeing their competitors and thinking about mergers. This taste for potential consolidation is even stretching beyond the industrial world — Imperial College and University College London are also contemplating joining forces.

What can one do to survive this merger madness? All the usual jobs advice applies. Update your CV. Network. Publish. Gain new skills...and wait for spring.

Paul Smaglik
Naturejobs editor



Contents

CAREERS AND RECRUITMENT

Computational neuroscience
finds its feet p4

Japan bridges the gap between
theory and experiment in
neuroscience p7

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

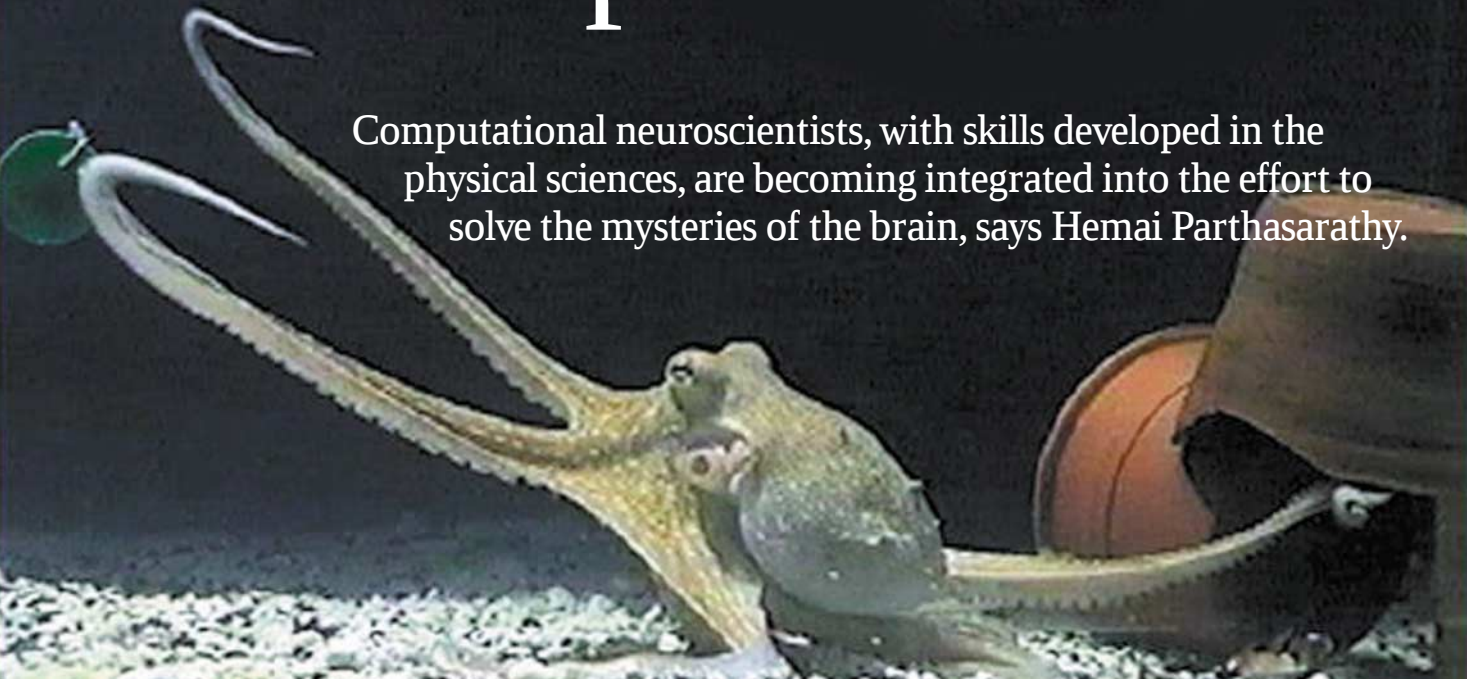
RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS

From scepticism to acceptance

Computational neuroscientists, with skills developed in the physical sciences, are becoming integrated into the effort to solve the mysteries of the brain, says Hemai Parthasarathy.



Open arms: how the octopus's brain controls movement is the focus of studies at the Center for Neural Computation in Jerusalem.

On course: Lyle Graham believes the training given in recent years has laid a strong foundation for the future of computational neuroscience.

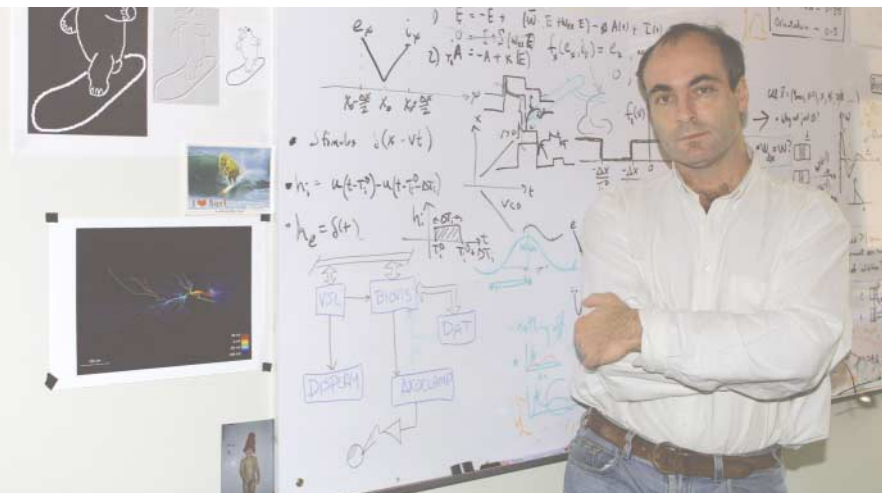
Modelling neuronal systems was once considered to be a marginal activity. But with more data and better computational techniques now available, opportunities for research and education are opening up. Terry Sejnowski, one of the discipline's current leaders, began to model neural processes as a secret side-project while he pursued his postgraduate course in astrophysics at Princeton University in New Jersey. Back then, in the mid-1970s, doing neural modelling in a physics department was "radical and unthinkable", says Robert Desimone, scientific director of the US National Institute of Mental Health in Bethesda,

Maryland. Over 20 years later, Sejnowski now runs the computational neuroscience laboratory at the Salk Institute for Biological Studies in La Jolla, California. And around the world physicists are playing an important part in transforming neurobiology.

But this shift in opinion took time. Neuroscientist Lyle Graham, who is now at the Laboratory for Computational Neuroscience in Gif-sur-Yvette, France, had similar experiences to Sejnowski. In the 1980s, while doing graduate work at the Massachusetts Institute of Technology, Graham began to do some neural modelling. The members of the electrical engineering faculty on his thesis committee gave him a hard time, he says. "Modelling had a huge reputation for being incredibly vague," says Graham. "I'm not sure I ever did convince them."

Even about 10 years ago, some biologists remained unconvinced about how useful or feasible modelling could be, and they viewed the approach with "hostility", says Charles Stevens, a neuroscience professor at the Salk Institute. But in the past few years more and more neurobiologists have begun to integrate computational and experimental approaches. They may be making this move because they now find themselves facing huge data sets and they need computationally sophisticated analysis — not only to handle the data but to design the questions.

With 60 billion neurons and 60 trillion synapses, the brain is arguably the most sophisticated computing device on the planet. And, of course, the brain is also a biological organ. As a result, physicists and engineers have had to join forces with biologists



What is neuroinformatics?

Although it requires strong computing skills, neuroinformatics is primarily about developing the tools and databases to aid neuroscientists manipulate, store and share their data. It is more closely related to engineering than to basic science. The newly launched journal *Neuroinformatics* advertises itself as the

“first journal devoted to the information-science infrastructure of neuroscience”. And next year, the Marine Biological Laboratory in Woods Hole, Massachusetts, is launching a two-week neuroinformatics summer course. With its own journal, workshops and funding initiatives, neuroinformatics is a field of its own. **H.P.**

to form a multipronged attack on the brain's secrets. A new breed of computational neuroscientist seems to be emerging that is trained specifically to study the brain with an interdisciplinary tool kit.

And, based on anecdotal evidence, the job market for computational neuroscientists seems to be quite healthy. Faculty and postdoc adverts in neuroscience often explicitly mention computational neuroscience, and many physics departments are also moving into biological physics, including some aspects of neuroscience.

Broadly speaking, computational neuroscientists come in three flavours. First and most obvious are those who make analytical or computational models of the nervous system, either at the level of single neurons or as neural networks. Another set devises tools to find the best ways to analyse experimental data. The third group informs experimental design by developing more sophisticated questions — creating stimulus sets to test neuronal response, for instance.

Physicists and mathematicians already have some basic skills in all three categories. But for those with strong quantitative skills and computer knowledge who want to get into neuroscience there is one major challenge — learning the biology, says Laurence

Abbott, a physicist-turned-computational neuroscientist at Brandeis University in Waltham, Massachusetts. “It’s like moving to a foreign country and learning a new language,” he says. “One day you wake up and find yourself part of the culture.”

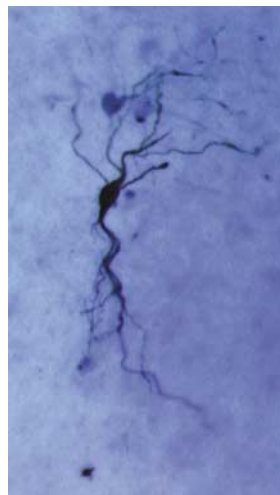
For Abbott, the secret was finding someone to help him learn that language and culture. Eve Marder, a neurophysiologist at Brandeis University acted as his guide and colleague in the foreign world of experimental biology.

One of the biggest obstacles in learning the language is the different way in which physicists and biologists approach scientific inquiry, notes Stevens. In physics, people know what the big questions are — the problem is finding a way to answer them. But biologists often don’t know what the question is.

A DIFFERENT APPROACH

Without a history of thinking about the biological problems involved, physical scientists may be tempted to use their analytical tools to address questions that they are not suited to. This is so common a mistake that the computational vision course held in the summer at Cold Spring Harbor Laboratory in New York doesn’t accept people straight out of a physics PhD, says Paul Glimcher, a primate neurophysiologist at New York University and one of the course’s organizers. He and his colleagues focus instead on educating experimental neurobiologists to think more quantitatively.

There is another way to make the switch — involve a collaborator with complementary expertise. Yuan Liu, a programme director at the US National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, emphasizes



Model behaviour: understanding how neurons work lies at the heart of computational neuroscience.

Learning the trade

Aspiring computational neuroscientists have some excellent options for learning or honing their skills outside their home institutions. The Marine Biological Laboratory in Woods Hole, Massachusetts, offers a four-week summer course on methods in computational neuroscience. Its counterpart in Europe, the European Union advanced course in computational neuroscience, is currently

based in Portugal. Both courses offer hands-on training in computational methods.

There are several shorter, more specialized courses and workshops, often based on specific software. Cold Spring Harbor Laboratory in New York also offers a course in computational neuroscience, but it concentrates on the visual system.

Robert Güti, a doctoral student at the University of

Freiburg, Germany, took the Woods Hole course in 2000. His background is in theoretical physics, although he was working in a computational neuroscience laboratory. Meeting Haim Sompolinsky, a professor at the Hebrew University of Jerusalem who was lecturing on the course, shifted the focus of his studies and their collaboration has continued for two years. “The environment that this course allows is very

good, says Güti. “You get the chance to really become friends and if there is a scientific match, you have everything you need to start a collaboration.”

Lyle Graham, a neuroscientist at the Laboratory for Computational Neuroscience in Gif-sur-Yvette, France, who taught on the Woods Hole course at its inception, echoes the view of many. “These courses have done a great job in building a generation,” he says. **H.P.**

Neuroscience and beyond

The route from the physical sciences into neurobiology doesn't necessarily stop there. Zoubin Ghahramani, who trained in computational neuroscience at the Massachusetts Institute of Technology, has recently moved into bioinformatics while working at the Gatsby

Computational Neuroscience Unit at University College London.

He notes that the switch has helped him to deal with many of the frustrations he encountered working in neuroscience, notably the gulf between the ability to develop a sophisticated

theory and the availability of experimental tools to test its relevance to biological systems.

He finds it satisfying that computational models in bioinformatics are clearly of practical importance. "A model of the genes interacting to produce the immune responses underlying

rheumatoid arthritis is obviously useful," he says. And he feels that the field is extremely receptive to computational methods.

"Who knows. Maybe the bubble will burst soon. But I really feel that there are still a huge number of open problems waiting for good technical solutions." **H.P.**

Causing a reaction: Rodney Douglas, director of the Institute of Neuroinformatics in Zurich, stimulates a processor designed to emulate visual processing in the brain (bottom). The institute's work made an appearance at the Swiss Expo. 02 in the form of 'Ada' an artificial creature that interacted with visitors using tools such as the intelligent floor (below).



that experimental neuroscientists who want help in designing sophisticated analytical tools should start collaborations before they start collecting the data. Once you have trained the rat, implanted electrodes in specific parts of its brain, and saved your data as averages, it is too late to realize that you could best test a theory by comparing activity in one brain region with another simultaneously, or by looking at responses over single trials.

To help bridge the divide between physical and biological sciences in neuroscience, a number of dedicated institutes have been created over the years. The five Sloan-Swartz Centers for Theoretical Neurobiology, set up in 1994, have set the standard for excellence in computational neuroscience. They are

based at Brandeis University, the California Institute of Technology, New York University, the Salk Institute, and the University of California, San Francisco.

Outside the United States, institutes with the goal of integrating theory and experiment have also emerged. In Zurich, the Institute of Neuroinformatics has fostered

an interdisciplinary approach since it was founded in 1995, and in Germany, computational neuroscience centres are being discussed at several universities to concentrate and enhance existing research activities. In France, the organizational structure of research does not easily accommodate such centres — instead, there are several research groups whose mission statement explicitly refers to computational neuroscience. And the Hebrew University of Jerusalem has established a strong centre for computational neuroscience at the graduate student level.

A NEW BREED

The doctoral programme at the Interdisciplinary Center for Neural Computation at the Hebrew University is 10 years old, and has students from biology, computer science, physics, mathematics and psychology. There are five compulsory courses given over three semesters that cover the basics — anatomy/physiology of the nervous system, neural networks, information theory, psychology and neuroscience methods.

The programme encourages theoreticians to consider problems of the nervous system. Training students in both theoretical and experimental approaches helps the graduates of the programme collaborate with others better — independently of which side of the fence they favour, says Eilon Vaadia, director of the centre's graduate programme.

Mayank Mehta, a physicist-turned-neuroscientist at the Massachusetts Institute of Technology, is encouraged by the new generation of computational neuroscientists who are comfortable in both worlds. At a recent visit to the University of California, San Diego, he was impressed that the students in the computational neuroscience programme had to rotate between experimental and theoretical labs. He asked every single student whether they had written a code in any language beyond Fortran, and whether they had dealt with live animals. "Every one of them had done both," he says.

But Abbott cautions that fluency in both languages is difficult to achieve. "It's pretty hard for students to learn both," he says. Computational neuroscientists are likely to continue to specialize in one side or the other. But another skill — communication — will be the key to making that specialism work. ■

Hemai Parthasarathy is a biological sciences editor at *Nature*.

T. DELBRUCK/INI

T. DELBRUCK/INI

A clash of two cultures

A lack of formal training is hampering Japan's efforts in computational neuroscience, says Robert Triendl.

In Japan, much like the rest of the world, uniting physics and biology in an effort to understand the brain has been a long process. But there are now signs of a strengthened resolve to combine experimental and theoretical approaches. Masao Ito, head of the RIKEN Brain Science Institute (BSI) in Saitama, sees an advantage in such a combination.

Cellular and molecular approaches can lead to new drugs, Ito says, but they do not help us to understand the brain as a complex system. But a purely computational route could miss potential applications. "I believe the merger of these two directions is what we should pursue in the coming years," Ito says.

Such links are now forming, says Mitsuo Kawato, a computational neuroscientist at the Advanced Telecommunications Research Institute International (ATR) in Kyoto. Research activities within Kawato's group combine experimental research in cognitive neuroscience with computer modelling and the development of robotic model systems. "Humanoid robots help us to verify our theories or to develop new hypotheses while building such devices," he says.

At the BSI, interactions between bench neuroscientists and computer scientists or mathematicians are also developing — albeit slowly. Differences in publication practices and research evaluation remain an important impediment to collaborations between the two camps, says Shun-ichi Amari, a mathematician who specializes in algorithms related to learning and cognition.

A COMMON APPROACH

Yoko Yamaguchi, director of the BSI's Laboratory for Dynamics of Emergent Intelligence, says that successful collaboration between experimentalists and theoreticians requires a shared set of concepts and research tools, such as an experimental model system that provides opportunities for both experimental research and theoretical analysis. But few experimental neuroscientists in Japan have been exposed to computational neuroscience during their training.



Researchers working with Mitsuo Kawato (below) build and run humanoid robots (left) to test their theories on cognitive neuroscience.

Cultural differences also provide an impediment. In Japan, most neuroscientists are educated in medical schools, whereas researchers specializing in such areas as neural networks or machine learning typically have backgrounds in electrical engineering or physics. Apart from a widely acclaimed summer school set up by the BSI, and a similarly successful event run by the Japanese Neural Network Society, there are no dedicated graduate-education programmes in computational neuroscience in Japan.

Employment opportunities for young researchers in computational neuroscience remain limited — there are few university positions, and research tends to be restricted to institutions such as the BSI, ATR or Sony Computer Sciences Laboratories in Tokyo.

WIRED UP FOR THE FUTURE

There are some promising developments at a few of the smaller private universities, such as the Kyushu Institute of Technology in Fukuoka, which established a department of brain sciences and engineering in 2000. But the overall job situation for young scientists in computational neuroscience is hardly encouraging.

There is some hope that the upcoming reform of national universities could provide opportunities to establish new, multidisciplinary graduate schools. But Kawato, who chairs a special committee at the education ministry on multidisciplinary research, says that it won't be easy for computational neuroscientists to find an institutional home in the universities. "In the present funding climate, the first question that government officials ask is always about the economic impact," he says.

Robert Triendl is a freelance science writer based in Tokyo.

RIKEN Brain Science Institute

♦ www.brain.riken.go.jp

Japanese Neural Network Society

♦ www.jnns.org/English

Advanced Telecommunications Research Institute International

♦ www.atr.co.jp/index-e2.html



Shun-ichi Amari (above) and Yoko Yamaguchi (below) want to link experiment and theory.

